University of California, Berkeley
U.C. Berkeley Division of Biostatistics Working Paper Series

Year 2010 *Paper* 259

Targeted Maximum Likelihood Based Causal Inference

Mark J. van der Laan*

*University of California - Berkeley, laan@berkeley.edu

Targeted Maximum Likelihood Based Causal Inference

Mark J. van der Laan

Abstract

Given causal graph assumptions, intervention-specific counterfactual distributions of the data can be defined by the so called G-computation formula, which is obtained by carrying out these interventions on the likelihood of the data factorized according to the causal graph. The obtained G-computation formula represents the counterfactual distribution the data would have had if this intervention would have been enforced on the system generating the data. A causal effect of interest can now be defined as some difference between these counterfactual distributions indexed by different interventions. For example, the interventions can represent static treatment regimens or individualized treatment rules that assign treatment in response to time-dependent covariates, and the causal effects could be defined in terms of features of the mean of the treatment-regimen specific counterfactual outcome of interest as a function of the corresponding treatment regimens. Such features could be defined nonparametrically in terms of so called (nonparametric) marginal structural models for static or individualized treatment rules, whose parameters can be thought of as (smooth) summary measures of differences between the treatment regimen specific counterfactual distributions.

In this article, we provide templates for implementation of the targeted maximum likelihood estimator of causal effects of multiple time point interventions. This involves the use of loss-based super-learning to obtain an initial estimate of the unknown factors of the G-computation formula, and subsequently, applying a target-parameter specific optimal fluctuation function (least favorable parametric submodel) to each estimated factor, estimating the fluctuation parameter(s) with maximum likelihood estimation, and iterating this updating step till convergence. The targeted maximum likelihood step makes the resulting estimator of the causal effect double robust in the sense that it is consistent if either the initial estimator is

consistent, or the estimator of the optimal fluctuation function is consistent. The optimal fluctuation function is correctly specified if the conditional distributions of the nodes in the causal graph one intervenes upon are correctly specified. The latter conditional distributions often comprise the so called treatment and censoring mechanism. Selection among different targeted maximum likelihood estimators (e.g., indexed by different initial estimators) can be based on loss-based cross-validation such as likelihood based cross-validation or cross-validation based on another appropriate loss function for the distribution of the data. Some specific loss functions are mentioned in this article.

In this article, a variety of interesting observations about targeted maximum likelihood estimation are made, and a concrete template for the practical implementation of targeted maximum likelihood estimation is presented as well. In addition, we demonstrate it for estimation of a causal effect of dynamic treatment rules defined in terms of a marginal structural working model, inspired by HIV applications.

1 Introduction.

The data structure on the experimental unit can often be viewed as a time-series in discrete time, possibly on a fine scale. At many time points nothing might be observed and at possibly irregular spaced time-points events occur and are measured, where some of these events occur at the same time. A specified ordering of all measured variables which respects this time-ordering and possibly additional knowledge about the ordering in which variables were realized, implies a graph in the sense that for each observed variable we can identify a set of parent nodes of that observed variable, defined as the set of variables occurring before the observed variable in the ordering. The likelihood of this unit specific data structure can be factorized accordingly in terms of the conditional distribution of a node in the graph, given the parents of that node, across all nodes. This particular factorization of the likelihood puts no restriction on the possible set of data generating distribution, but the ordering affects the so called G-computation formula for counterfactual distributions of the data under certain interventions implied by this ordering. Beyond the factorization of the likelihood in terms of a product of conditional distributions, the G-computation formula involves specifying a set of nodes in the time-series/graph as the variables to intervene upon, and specifying the intervention for these nodes. These interventions could be rules that assign the value for the intervention node (possibly) in response to the observed data on the (observed) parents of the intervention node. The G-computation formula is now defined as the product, across all nodes, excluding the intervention nodes, of the conditional distribution of a node, given the parent nodes with the intervention nodes in the parent set following their assigned values.

If it is known that the conditional distribution of a node only depends on a subset of the parents that were implied by the ordering, then that knowledge should be incorporated by reducing the parent set to its correct set. This kind of knowledge does reduce the size of the model for the data generating distribution (and such assumptions can indeed be tested from the data).

The G-computation formula provides a probability distribution of the intervention specific data structure. Under certain causal graph conditions on a causal graph on an augmented set of nodes which includes unobserved nodes beyond the observed nodes (Pearl (2000)), such as no unblocked back-door path from intervention node to future/downstream nodes, for each intervention node, this G-computation formula equals the counterfactual distribution

of the data structure if one would have enforced the specified intervention on the system described by the causal graph.

We remind the reader that a causal graph on a set of nodes states that each node is a deterministic function of its parents. It typically represents a set of so called causal assumptions that cannot be learned from the data. Given a declared causal graph on a set of nodes, one can formally state what assumptions on this causal graph are needed in order to claim that a specified G-computation formula for the *observed* nodes corresponds with the G-computation formula for the causal graph on the full set of nodes (that includes the unobserved nodes), where the latter G-computation formula is then viewed as the gold-standard representing the causal effect of interest (Pearl (2000)).

Either way, the time-ordering and possible known additional known ordering does provide a statistical graph for the data as explained above, and a corresponding G-computation formula.

In this article we are concerned with (semi-parametric) efficient estimation of the "causal" effects viewed as parameters of the G-computation formula based on observing n independent and identically observations $O_1, \ldots, O_n$ of O. Specifically, we are concerned with estimation of parameters of the G-computation formula implied by a particular statistical graph on the observed data structure O, in the semiparametric model that makes no assumptions about each node-specific conditional distribution in the graph, given its parents.

Formally, the density of O is modeled as

$$p_0(O) = \prod_j P(N(j) \mid Pa(N(j))$$

where $N(j)$ denote the nodes in the graph representing the observed variables, $Pa(N(j))$ denote the parents of $N(j)$, and we make no assumptions on each conditional distribution of $N(j)$, beyond that $N(j)$ only depends on $Pa(N(j))$. Note, however, as remarked above, if the parent sets induce more structure than parent sets implied by an ordering of all observed variables, then this statistical graph of p_0 might implies a real (i.e., not just nonparametric) semiparametric model on p_0, corresponding with a variety of conditional independence assumptions.

Even if the (non-testable) causal assumptions required to interpret the G-computation formula as a counterfactual distribution on a system fail to

hold, assuming that the ordering of the likelihood respects the ordering w.r.t. the intervention nodes (i.e., it correctly states what variables are pre or post intervention for each intervention node), the target parameters often still represent effects of interest aiming to get as close to a causal effect as the data allows. In particular, one can simply interpret the G-computation parameters for what they are, namely well defined effects of interventions on the distribution of the data: see van der Laan (2006) for more discussion of the role of causal parameters in variable importance analysis.

It is important to note that the probability density p_0 of the observed data structure O, factored by the statistical graph, can be represented as a product of two factors, the first factor Q_0 that identifies the G-computation formulas for interventions, and the second factor g_0 representing product over the intervention nodes of the conditional distribution of the intervention nodes: $p_0 = Q_0 g_0$. We often refer to the second factor as the censoring and/or treatment mechanism in case the intervention nodes correspond with censoring variables and/or treatment assignments. We will denote the true probability distribution of the data-structure on the experimental unit with P_0, and its probability density with p_0.

A variety of estimators of causal effects of multiple time-point interventions, including handling censored data (by, enforcing no-censoring as part of the intervention) have been proposed: Inverse Probability of Censoring Weighted (IPCW) estimators, Augmented IPCW-estimators (which are double robust), maximum likelihood based estimators, and targeted Maximum Likelihood Estimators (which are double robust). The IPCW and augmented-IPCW estimators fall in the category of estimating equation methodology (van der Laan and Robins (2003)). The augmented-IPCW estimator is defined as a solution of an estimating equation in the target parameter implied by the so called efficient influence curve. Maximum likelihood based estimators involve estimation of the distribution of the data and subsequent evaluation of the target parameter. Traditional maximum likelihood estimators are not targeted towards the target parameter, and are thereby, in particular, not double robust.

Targeted maximum likelihood estimators (T-MLE) are two stage estimators, the first stage applies regularized maximum likelihood based estimation, where we advocate the use of loss-based super-learning to maximize adaptivity to the true distribution/G-computation formula of data (van der Laan et al. (2007)), and the second stage targets the obtained fit from the first stage towards the target parameter of interest through a targeted maximum

likelihood step. This targeted maximum likelihood step removes bias for the target parameter if the censoring/treatment mechanism used in the targeted MLE step is estimated consistently. In this targeted maximum likelihood step the initial (first stage) estimator is treated as an off-set, and it involves the application of a fluctuation function to the offset, where the set of possible fluctuations represents a parametric model consisting of fluctuated versions of the offset. This parametric model is a so called least favorable parametric model in the sense that its maximum likelihood estimator listens as much to the data w.r.t. fitting the target parameter as a semiparametric model efficient estimator. Formally, it is the parametric submodel through the first stage estimator with the worst Cramer-Rao lower bound for estimation of the target parameter (at zero fluctuation), among all parametric submodels. (This worst case Cramer-Rao lower bound as achieved by this least favorable model is actually the semiparametric information bound defined as the variance of the efficient influence curve.) Given this least-favorable submodel, maximum likelihood estimation is used to fit the finite dimensional fluctuation parameter. Due to this parametric targeted maximum likelihood step the targeted maximum likelihood estimator is also double robust: the estimator is consistent if the initial first-stage estimator of the G-computation factor of the likelihood is consistent, or if the conditional distributions of the intervention nodes (i.e., censoring/treatment mechanism) are estimated consistently (as required to identify the fluctuation function used in targeted maximum likelihood step). In addition, under regularity conditions, the targeted MLE is (semiparametric) efficient if the initial estimator is consistent, and consistent and asymptotically linear if either the initial estimator or the treatment/censoring mechanism estimator is consistent.

Even though the augmented IPCW-estimator is also tailored to be double robust and locally efficient, targeted maximum likelihood estimation has the following important advantages relative to estimating equation methods such as the augmented-IPCW estimator: 1) the T-MLE is a substitution estimator and thereby, contrary to the augmented IPCW-estimator, respects global constraints of the model such as that one might be estimating a probability in $[0, 1]$, 2) since, given an initial estimator, the targeted MLE step involves maximizing the likelihood along a smooth parametric submodel, contrary to the augmented IPCW-estimator, it does *not* suffer from multiple solutions of a (possibly non-smooth in the parameter) estimating equation, 3) contrary to the augmented IPCW-estimator, the T-MLE does *not* require that the efficient influence curve can be represented as an estimating function in the

target parameter, and thereby applies to *all* path-wise differentiable parameters, 4) it can use the cross-validated log-likelihood (of the targeted maximum likelihood estimator), or any other cross-validated risk of an appropriate loss function for the relevant factor Q_0 of the density (i.e., the G-computation formula) of the data, as principle criterion to select among different targeted maximum likelihood estimators indexed by different initial estimators or different choices of fluctuation models. The latter allows fine tuning of the initial estimator of Q_0 as well as the fine tuning of the estimation of the unknowns (e.g., censoring/treatment mechanism g_0) of the fluctuation function applied in the targeted maximum likelihood step, thereby utilizing the excellent theoretical and practical properties of the loss-function specific cross-validation selector. On the other hand, the augmented-IPCW estimator cannot be evaluated based on a loss function for Q_0 alone, but also requires a choice of loss function for g_0. The latter point 4) also allows the targeted MLE to be generalized to loss-based estimation of infinite dimensional parameters that can be approximated by pathwise differentiable parameters.

These important theoretical advantages have a substantial practical impact, by allowing one to construct estimators in a wider variety of applications, and with better finite sample and asymptotic mean squared error w.r.t. the target. This inspired us to implement targeted maximum likelihood estimation of causal effects of single time point treatment in a variety of data analyses, allowing for right-censoring of the time-till-event clinical outcome, and missingness of the clinical outcome. Even though we discussed the overall targeted maximum likelihood estimator for causal effect estimation of multiple time point interventions in technical reports (see van der Laan (2008)), in this article we aim to dive deeper into this challenge. In particular, our goal is to present templates that can be implemented with standard statistical software, and aim to understand the choices to be made. In future papers we will be implementing these methods on real and simulated data sets and use this paper as guidance.

The organization of this paper is as follows. Firstly, in Section 2 we start out with presenting the targeted MLE for sequentially randomized controlled trials. A specific targeted loss function is proposed to select among different targeted MLE indexed by different initial estimators, which results in maximally asymptotically efficient targeted MLE's (Rubin and van der Laan (2008)). Due to the double robustness of the targeted MLE this estimator is guaranteed to estimate the causal effect of interest consistently, so that confidence intervals and type-I error control are asymptotically valid. In addition,

the T-MLE utilizes all the data (including time-dependent biomarkers) and thereby has great potential for large efficiency gains and bias reductions in these sequentially randomized controlled trials.

In Section 3 we develop and present a general targeted MLE for any time-series data structure, applicable to sequentially randomized controlled trials with censoring and missingness, as well as longitudinal observational studies. The integration of loss-based (super) learning to build and select among targeted MLE's is made explicit again, and targeted loss functions are proposed for that purpose. In Section 4 we present a detailed and concrete template for the targeted MLE. We also discuss the natural extension to collaborative targeted MLE, which involves selection among different fluctuation functions indexed by different censoring/treatment mechanisms in order to fine tune the targeted maximum likelihood step for effective (i.e., no need to focus on bias reduction that has already been taken care off by initial estimator) bias reduction, as presented in van der Laan and Gruber (2009). In Section 5 we present additional concrete details by describing the targeted MLE for estimation of causal effects defined by a marginal structural model for individualized treatment rules in an HIV application. In Section 6 we consider the targeted MLE for a causal effect of a point treatment on future outcome, incorporating time-dependent covariates and allowing for right-censoring. We end with a discussion in Section 7.

Some overview of relevant literature

The construction of efficient estimators of path-wise differentiable parameters in semi-parametric models requires utilizing the so called efficient influence curve defined as the canonical gradient of the path-wise derivative of the parameter. This is no surprise since a fundamental result of the efficiency theory is that a regular estimator is efficient if and only if it is asymptotically linear with influence curve equal to the efficient influence curve. We refer to Bickel et al. (1997), and Andersen et al. (1993). There are two distinct approaches for construction of efficient (or locally efficient) estimators: the estimating equation approach that uses the efficient influence curve as an estimating equation (e.g., one-step estimators based on the Newton-Raphson algorithm in Bickel et al. (1997)), and the targeted MLE that uses the efficient influence curve to define a targeted fluctuation function of an initial estimator, and maximizes the likelihood in that targeted direction.

The construction of locally efficient estimators in censored data models in

which the censoring mechanism satisfies the so called coarsening at random assumption (Heitjan and Rubin (1991), Jacobsen and Keiding (1995), Gill et al. (1997)) has been a particular focus area. This includes also the theory for locally efficient estimation of causal effects, since the causal inference data structure can be viewed as a missing data structure on the intervention-specific counterfactuals, and the sequential randomization assumption (SRA) implies the coarsening at random assumption on the missingness mechanism, while SRA still does not imply any restriction on the data generating distribution. A particular construction of counterfactuals from the observed data structure, so that the observed data structure augmented with the counterfactuals satisfies the consistency (missing data structure) and sequential randomization assumption, is provided in Yu and van der Laan (2002), providing an alternative to the implicit construction presented earlier in Gill and Robins (2001), thereby showing that, without loss of generality, one can view causal inference as a missing data structure estimation problem: the importance of the causal graph is that it makes explicit the definition of the counterfactuals of interest (i.e., full data in the censored data model).

The theory for inverse probability of censoring weighted estimation and the augmented locally efficient IPCW estimator based on estimating functions defined in terms of the orthogonal complement of the nuisance tangent space in CAR-censored data models (including the optimal estimating function implied by efficient influence curve) was originally developed in Robins (1993), Robins and Rotnitzky (1992). Many papers have been building on this framework (see van der Laan and Robins (2003) for a unified treatment of this estimating equation methodology and references). In particular, double robust locally efficient augmented IPCW-estimators have been developed (Robins and Rotnitzky (2001b), Robins and Rotnitzky (2001a), Robins et al. (2000b), Robins (2000a), van der Laan and Robins (2003), Neugebauer and van der Laan (2005), Yu and van der Laan (2003)).

Causal inference for multiple time-point interventions under sequential randomization started out with papers by Robins in the eighties: e.g. Robins (1986), Robins (1989). The popular propensity score methods to assess causal effects of single time point interventions (e.g., Rosenbaum and Rubin (1983), Sekhon (2008), Rubin (2006)) are not double robust (i.e., rely on correct specification of propensity score), have no natural generalization to multiple time-point interventions, and are also inefficient estimators for single time point interventions (Abadie and Imbens (2006)), relative to the locally efficient double robust estimators such as the augmented IPCW estimator, and

the targeted MLE.

Structural nested models and marginal structural models for static treatments were proposed by Robins as well: Robins (1997b), Robins (1997a), Robins (2000b). Many application papers on marginal structural models exist, involving the application of estimating equation methodology (IPCW and DR-IPCW): e.g., Hernan et al. (2000), Robins et al. (2000a), Bryan et al. (2003), Yu and van der Laan (2003). In van der Laan et al. (2005) history adjusted marginal structural models were proposed as a natural extension of marginal structural models, and it was shown that the latter also imply an individualized treatment rule of interest (a so called history adjusted statically optimal treatment regimen): see Petersen et al. (2005) for an application to the when to switch question in HIV research.

Murphy et al. (2001) present a nonparametric estimator for a mean under a dynamic treatment in an observational study. Structural nested models for modeling and estimating an optimal dynamic treatment were proposed by Murphy (2003), Robins (2003), Robins (2005a), Robins (2005b). Marginal structural models for a user supplied set of dynamic treatment regimens were developed and proposed in van der Laan (2006), van der Laan and Petersen (2007) and, simultaneously and independently, in Robins et al. (2008). van der Laan and Petersen (2007) also includes a data analysis application of these models to assess the mean outcome under a rule that switches treatment when CD4 count drops below a cut-off, and the optimal cut-off is estimated as well. Another practical illustration in sequentially randomized trials of these marginal structural models for realistic individualized treatment rules is presented in Bembom and van der Laan (2007).

Unified loss-based learning based on cross-validation was developed invan der Laan and Dudoit (2003), including construction of adaptive minimax estimators for infinite dimensional parameters of the full data distribution in CAR-censored data and causal inference models: see also van der Laan et al. (2006), van der Vaart et al. (2006), van der Laan et al. (2004), Dudoit and van der Laan (2005), Keleş et al. (2002), Sinisi and van der Laan (2004). This research establishes, in particular, finite sample oracle inequalities, which state that the expectation of the loss-function specific dissimilarity between the the cross-validated selected estimator among the library of candidate estimators (trained on training samples) and the truth is smaller or equal than the expectation of the loss-function specific dissimilarity between the best possible selected estimator and the truth plus a term that is bounded by a constant times the logarithm of the number of candidate

estimators in the library divided by the sample size. The only assumption this oracle inequality relies upon is that the loss function is uniformly bounded. These oracle results for the cross-validation selector inspired a unified super-learning methodology. This methodology first constructs a set of candidate estimators, proposes a family of weighted combinations of these candidate estimators indexed by a weight vector, and uses cross-validation to determine a weighted combination with optimal cross-validated risk. Under the assumption that the loss function is uniformly bounded, and the number of estimators is polynomial in sample size, the resulting estimator (super learner) is either asymptotically equivalent with the oracle selected estimator among the library of weighted combinations of the estimators, or it achieves the optimal parametric rate of convergence (i.e. one of estimators corresponds with correctly specified parametric model) up till (worst case) log-n-factor. We refer to van der Laan et al. (2007), Polley and van der Laan (2009).

The super-learning methodology applied to a loss function for the G-computation formula factor, Q_0, of the observed data distribution, provides substitution estimators of ψ_0. However, although these super learners of Q_0 are optimal w.r.t. the dissimilarity with Q_0 implied by the loss function, the corresponding substitution estimators will be overly biased for a smooth parameter mapping Ψ. This is due to the fact that cross-validation makes optimal choices w.r.t. the (global) loss-function specific dissimilarity, but the variance of $\Psi(Q_n)$ is of smaller order than the variance of Q_n itself.

van der Laan and Rubin (2006) integrates the loss-based learning of Q_0 into the locally efficient estimation of pathwise differentiable parameters, by enforcing the restriction in the loss-based learning that each candidate estimator of Q_0 needs to be a targeted maximum likelihood estimator (thereby, in particular, enforcing each candidate estimator of Q_0 to solve the efficient influence curve estimating equation). Another way to think about this is that each loss function $L(Q)$ for Q_0 has a corresponding targeted loss function $L(Q^*)$, Q^* representing the targeted maximum likelihood estimator applied to initial Q, and we apply the loss-based learning to the latter targeted version of the loss function $L(Q)$. Rubin and van der Laan (2008) propose the square of efficient influence curve as a valid and sensible loss function $L(Q)$ for selection and estimation of Q_0 in models in which g_0 can be estimated consistently, such as in randomized controlled trials.

The implications of this targeted loss-based learning are that Q_0 is estimated optimally (maximally adaptive to the true Q_0) w.r.t. the targeted

loss function $L(Q^*)$ using the super-learning methodology, and due to the targeted MLE step the resulting substitution estimator of ψ_0 is now asymptotically linear as well if the targeted fluctuation function is estimated at a good enough rate (and only requiring adjustment by confounders not yet accounted for by initial estimator: see collaborative targeted MLE): either way, asymptotic bias reduction for the target parameter will occur as long as the censoring/treatment mechanism is estimated consistently. In addition, since the targeted maximum likelihood step involves additional maximum likelihood fitting, generally speaking, no loss in bias will occur, even if the wished fluctuation function is heavily misspecified.

Targeted MLE have been applied in a variety of estimation problems: Bembom et al. (2008), Bembom et al. (2009) (physical activity), Tuglus and van der Laan (2008) (biomarker analysis), Rosenblum et al. (2009) (AIDS), van der Laan (2008) (case control studies), Rose and van der Laan (2008) (case control studies), Rose and van der Laan (2009) (matched case control studies), Moore and van der Laan (2009) (causal effect on time till event, allowing for right-censoring), van der Laan (2008) (adaptive designs, and multiple time point interventions), Moore and van der Laan (2007) (randomized trials with binary outcome). We refer to van der Laan et al. (September, 2009) for collective readings on targeted maximum likelihood estimation.

In van der Laan and Gruber (2009) we use the loss-based cross-validation to not only select among different initial estimators for the targeted maximum likelihood estimators, but it is also used to select the fit of the fluctuation function applied to the initial estimator (and thus the fit of the censoring and treatment mechanism). This results in a so called collaborative double robust targeted maximum likelihood estimator, which utilizes the collaborative double robustness of the efficient influence curve, which is stronger robustness result than the regular double robustness of the efficient influence curve the double robust estimators rely upon. These collaborative double robust estimators select confounders for the censoring and treatment mechanism in response to the outcome and the initial estimator of Q_0, thereby allowing for more effective bias reductions by the resulting fluctuation functions (as predicted by theory and observed in practice). Simulations and data analysis illustrating the excellent performance of the collaborative double robust T-MLE are presented in van der Laan and Gruber (2009).

2 The T-MLE in multi-stage sequentially randomized controlled trials

Consider a sequentially randomized trial in which one randomly samples a patient from a population, one collects at baseline covariates $L(0)$, and one randomizes the patient to a first line treatment $A(0)$. Subsequently, one collects an intermediate biomarker $L(1)$, and based on this intermediate clinical response one randomizes the patient to a second line treatment $A(1)$. Finally, one collects the clinical outcome Y of interest at a fixed point in time. This is experiment is carried out for n patients.

We first discuss two of such sequentially randomized cancer trials.

Anderson Cancer Center Prostate Cancer Two Stage Trial: Thall et al. (2000) present an analysis of the first clinical trial in oncology that makes use of sequential randomization. During this trial, prostate cancer patients who were found to be responding poorly to their initially randomly assigned regimen (among four treatments) were re-randomized to the remaining three candidate regimens. The clinical outcomes of interest was response to treatment at a particular point in time or time till death. In contrast to conventional trials based on a single randomization, this design allows the investigator to study adaptive treatment strategies that adjust a patients treatment in response to the observed course of the illness. Such adaptive strategies, also referred to as dynamic or individualized treatment rules, form the basis of common medical practice in cancer chemotherapy, with physicians typically facing the following questions: Which regimen should be used to initially treat a patient? Which regimen should the patient be switched to if the front-line regimen fails? Given an observed intermediate outcome such as a change in tumor size or PSA level, what threshold should be used to decide that the current regimen is failing? In recent years, sequentially randomized trials have been recognized as being uniquely suited to the study of these exciting questions (Thall et al. (2000), Lavori and Dawson (2000), Lavori and Dawson (2004), Murphy (2005)), with researchers in other clinical areas also beginning to implement this design (Rush et al. (2003), Schneider et al. (2006), Swartz et al. (2007)). The original results of Thall et al. (2000) focus on fitting logistic regression models for the different stage-specific factors of the likelihood. We can apply the T-MLE to estimate the mean outcomes under the 12 dynamic treatment rules indexed by first line therapy and the second line switching therapy, and also incorporate the handling of

the right-censoring.

E4494 Eastern Oncology Trial: Another example is the cancer trial E4494, a lymphoma study of rituximab therapy that had both induction and maintenance rituximab randomizations, where the second randomization of maintenance versus observation was based on intermediate response to the initial treatment. The clinical outcome of interest was time till death.

Let's denote the observed data structure on a randomly sampled patient from the target population with $O = (L(0), A(0), L(1), A(1), Y = L(2))$. For simplicity and the sake of presentation, we will assume that $A(0), L(1)$ and $A(1)$ are binary.

The likelihood can be factorized as

$$p(O) = \prod_{j=0}^{2} P(L(j) \mid \bar{L}(j-1), \bar{A}(j-1)) \prod_{j=0}^{1} P(A(j) \mid \bar{A}(j-1), \bar{L}(j)),$$

where the first factors will be denoted with $Q_{L(j)}$, $j = 0, 1, 2$, and the latter factors denote the treatment mechanism and are denoted with $g_{A(j)}$, $j = 0, 1$. We make the convention that for $j = 0$, $\bar{A}(j-1)$ and $\bar{L}(j-1)$ are empty. In a sequentially randomized controlled trial, the treatment assignment mechanisms $g_{A(j)}$, $j = 0, 1$, are known.

Suppose our parameter of interest is the treatment specific mean EY_d for a certain treatment rule d that assigns treatment $d_0(L(0))$ at time 0 and treatment $d_1(\bar{L}(1), A(0))$ at time 1. For example, $d_0(L(0)) = 1$ is a static treatment assignment, and $d_1(\bar{L}(1), A(0)) = I(L(1) = 1)1 + I(L(1) = 0)0$ assigns treatment 1 if the patients responds well to the first line treatment 1, and treatment 0 if the patients does not respond well to the first line treatment 1. We note that any treatment rule can be viewed as a function of $\bar{L} = (L(0), L(1))$ only, and therefore we will use the shorter notation $d(\bar{L}) = (d_0(L(0)), d_1(\bar{L}))$ for the two rules at times 0 and 1.

Note that $EY_d = \Psi(Q)$ for a well defined mapping Ψ. Specifically, we have $\Psi(Q) = E_{P_d} Y$, where the intervened distribution P_d of $(L(0), L(1), L(2))$ is defined by the G-computation formula:

$$P_d(\bar{L}) = \prod_{j=0}^{2} Q_{L(j),d}(\bar{L}(j)),$$

where, for notational convenience, especially in view of our representation for general data structures, we used the notation $Q_{L(j),d}(\bar{L}(j)) = Q_{L(j)}(L(j) \mid \bar{L}(j-1), \bar{A}(j-1) = d(\bar{L}(j-1)))$.

Instead of computing an analytic mean under P_d, this mean can also be approximated by simulating a large number of observations from this distribution and taking the mean of its last component $L(2)$. Note that P_d corresponds with simulating sequentially from the conditional distributions $Q_{L(0),d}, Q_{L(1),d}, Q_{L(2),d}$, for $L(0)$, $L(1)$, $L(2)$, respectively. Alternatively, this mean is calculated analytically as follows:

$$\begin{aligned}\Psi(Q) &= \sum_{l(0),l(1),y} y P_d(l(0),l(1),y) \\ &= \sum_y y \sum_{l(0),l(1)} P_d(l(0),l(1),y) \\ &= \sum_y y \sum_{l(0),l(1)} Q_{L(0)}(l(0)) Q_{L(1),d}(l(0),l(1)) Q_{Y,d}(l(0),l(1),y).\end{aligned}$$

If $Q_{L(0)}$ is replaced by the empirical distribution, then this reduces to

$$\Psi(Q) = \frac{1}{n}\sum_{i=1}^n \sum_y y \sum_{l(1)} Q_{L(1),d}(L_i(0),l(1)) Q_{Y,d}(L_i(0),l(1),y).$$

From this analytic expression it also follows that, even if Y is continuous, $\Psi(Q)$ only depends on the conditional distribution of Y through its mean. In this case of a 2-stage sequentially randomized controlled trial, the analytic evaluation of $\Psi(Q)$ seems preferable since it will be very fast to compute.

With this precise definition of the parameter as a mapping from the conditional distributions $Q_{L(j)}$, $j = 0, 1, 2$, to the real line, given an estimator Q_n, we obtain a substitution estimator $\Psi(Q_n)$ of ψ.

The targeted maximum likelihood estimator involves first defining an initial estimator of Q, and then a subsequent targeted maximum likelihood step according to a fluctuation function applied to this initial estimator, where this step is tailored to remove bias from the initial estimator for the purpose of estimating the parameter of interest ψ. The fluctuation function equals the least favorable parametric model through Q which is defined as the parametric submodel through Q which makes estimation of $\Psi(Q)$ hardest in the sense that the parametric Cramer-Rao Lower bound for the variance of an unbiased estimator is maximal among all parametric submodels. Intuitively, this is the parametric submodel for which the maximum likelihood estimator listens as much to the data w.r.t. fitting the target parameter as an efficient estimator in the large semiparametric model, and thereby can expected to

provide important bias reduction. This definition of least favorable model implies that a least favorable parametric model is a model that has a score at zero fluctuation equal to the efficient influence curve/canonical gradient of the pathwise derivative of the target parameter Ψ.

We use the following Theorem that provides the representation of the efficient influence curve which is needed to define the fluctuation function. This Theorem also provides the formula for the efficient influence curve for other parameters and for higher (than two) stage RCT's.

Theorem 1 *The efficient influence curve for $\psi = EY_d$ at the true distribution P_0 of O can be represented as*

$$D^* = \Pi(D_{IPCW} \mid T_Q),$$

where

$$D_{IPWC}(O) = \frac{I(\bar{A} = d(\bar{L}))}{g(d(\bar{L}) \mid X)} Y - \psi,$$

T_Q is the tangent space of Q in the nonparametric model, and Π denotes the projection operator onto T_Q in the Hilbert space $L_0^2(P_0)$ of square P_0-integrable functions of O, endowed with inner product $\langle h_1, h_2 \rangle = E_{P_0} h_1 h_2(O)$.

We have that this subspace

$$T_Q = \sum_{j=0}^{2} T_{Q_{L(j)}}$$

is the orthogonal sum of the tangent spaces $T_{Q_{L(j)}}$ of the $Q_{L(j)}$-factors, which consists of functions of $L(j), Pa(L(j))$ with conditional mean zero, given the parents $Pa(L(j))$ of $L(j)$, $j = 0, 1, 2$. Recall that we also denote $L(2)$ with Y. Let

$$D_j^* = \Pi(D^* \mid T_{Q_{L(j)}}),\ j = 0, 1, 2.$$

We have

$$\begin{aligned} D_0^*(O) &= E(Y_d \mid L(0)) - \psi \\ D_1^*(O)) &= \frac{I(A(0) = d_0(L(0)))}{g(d_0(L(0)) \mid X)} \times \\ &\{E(Y_d \mid L(0), A(0), L(1) = 1) - E(Y_d \mid L(0), A(0), L(1) = 0)\} \{L(1) - E(L(1) \mid L(0), A(0))\} \\ D_2^*(O) &= \frac{I(\bar{A} = d(\bar{L})}{g(\bar{A} \mid X)} \left\{L(2) - E(L(2) \mid \bar{L}(1), \bar{A}(2))\right\}. \end{aligned}$$

We note that

$$E(Y_d \mid L(0), A(0) = d_0(L(0)), L(1)) = E(Y \mid \bar{L}(1), \bar{A} = d(\bar{L})).$$

For general data structure $O = L(0), A(0), \ldots, L(K), A(K), Y = L(K+1)$*, and* $D_{IPCW} = D_1(O)/g(\bar{A} \mid X)$ *for some* D_1*, we have*

$$\begin{aligned}
&\Pi(D_{IPCW} \mid T_{Q_{L(j)}}) = \tfrac{1}{g(\bar{A}(j-1)|X)} \\
&\times \left\{ E\left(\textstyle\sum_{\bar{a}(j,K)} D_1(\bar{a}(j,K)) \mid \bar{L}(j), \bar{A}(j-1) \right) \right. \\
&\left. -E\left(\textstyle\sum_{\bar{a}(j,K)} D_1(\bar{a}(j,K)) \mid \bar{L}(j-1), \bar{A}(j-1) \right) \right\} \\
&= \tfrac{1}{g(\bar{A}(j-1)|X)} C_{L(j)}(\bar{L}(j-1), \bar{A}(j-1))(L(j) - E(L(j) \mid \bar{L}(j-1), \bar{A}(j-1)),
\end{aligned}$$

where

$$\begin{aligned}
C_{L(j)} \;=\; & E\left(\sum_{\bar{a}(j,K)} D_1(\bar{a}(j,K)) \mid L(j) = 1, \bar{L}(j-1), \bar{A}(j-1) \right) \\
& -E\left(\sum_{\bar{a}(j,K)} D_1(\bar{a}(j,K)) \mid L(j) = 0, \bar{L}(j-1), \bar{A}(j-1) \right).
\end{aligned}$$

Here we use the short-hand notation $\bar{a}(j,K) \equiv (a(j), \ldots, a(K))$ *and* $D_1(\bar{a}(j,K)) = D_1(\bar{A}(j-1), \bar{a}(j,K), \bar{L}_{\bar{a}(j,K)}(K+1))$*.*

This Theorem allows us now to specify the targeted maximum likelihood estimator.

The targeted maximum likelihood estimator: Consider now an initial estimator $Q_{L(j)n}$ of each $Q_{L(j)}$, $j = 0, 1, 2$. We will estimate the first marginal probability distribution $Q_{L(0)}$ of $L(0)$ with the empirical distribution of $L_i(0)$, $i = 1, \ldots, n$. We can estimate the conditional distributions of the binary $L(1)$ and the conditional mean of $Y = L(2)$ with machine learning algorithms (using logistic link for $Q_{L(1)}$, and, if Y is binary, also for Q_Y) such as the super learner represented by a data adaptively (based on cross-validation) determined weighted combination of a user supplied set of candidate machine learning algorithms estimating the particular conditional probability. We will now define fluctuations of this initial estimator $Q_{L(1)n} = Q_{L(1)n}(P_n)$ and $Q_{L(2)n} = Q_{L(2)}(P_n)$ which are particular

functions of the empirical probability distribution P_n. We will use notation $Q_n = (Q_{L(1)n}, Q_{L(2)n})$. Firstly, let

$$\text{Logit}Q_{L(1)n}(\epsilon) = \text{Logit}Q_{L(1)n} + \epsilon C_{L(1)}(Q_n, g_n)$$

be the fluctuation function of $Q_{L(1)n}$ with fluctuation parameter ϵ, where we added the covariate $C_{L(1)}(Q, g)$ defined as

$$\frac{I(A(0) = d_0(L(0))}{g_{A(0}(d_0(L(0)) \mid X)} \left\{E_Q(Y_d \mid L(0), A(0), L(1) = 1) - E_Q(Y_d \mid L(0), A(0), L(1) = 0)\right\}.$$

We refer to these covariate choices as *clever* covariates, since they represent a covariate choice that identifies a least favorable fluctuation model, thereby providing the wished targeted bias reduction. Similarly, if $Y = L(2)$ is binary, then let

$$\text{Logit}Q_{L(2)n}(\epsilon) = \text{Logit}Q_{L(2)n} + \epsilon C_{L(2)}(Q_n, g_n),$$

where the added the clever covariate

$$C_{L(2)}(Q, g)(\bar{L}(1), \bar{A}(1)) = \frac{I(\bar{A} = d(\bar{L})}{g(\bar{A} \mid X)}.$$

If Y is continuous, then we use as fluctuation model the normal densities with mean $E_{Q_{Y_n}}(Y \mid Pa(Y)) + \epsilon C_{L(2)}(Q_n, g_n)$, and constant variance σ^2, so that the MLE of ϵ is the linear least squares estimator, and the score of ϵ at $\epsilon = 0$ is $C_{L(2)}(Y - E_Q(Y \mid Pa(Y)))$, as required. We note that the above fluctuation function indeed satisfies that the score of ϵ at $\epsilon = 0$ equals the efficient influence curve $D^*(Q_n, g_n)$ as presented in the Theorem above.

One now estimates ϵ with the MLE.

$$\epsilon_n = \arg\max_{\epsilon} \prod_{j=1}^{2} \prod_{i=1}^{n} Q_{L(j)n}(\epsilon)(O_i).$$

One could also obtain a separate MLE of ϵ for each factor $j = 1, 2$. This process is now iterated till convergence, which defines the targeted MLE (Q_n^*, g_n), starting at initial estimator (Q_n, g_n), which does not involve updating of g_n.

We note that the ϵ_n for each factor separately can be estimated with standard logistic regression or linear regression software using as off-set the logit of the initial estimator and having a single clever covariate $C_{L(j)}(Q, g)$, $j = 1, 2$. If Y is also binary, the single/common ϵ_n defined above requires

applying a single logistic regression applied to repeated measures data set with one line of data for each of two factors, creating a clever covariate column that alternates the clever covariates $C_{L(1)}$ and $C_{L(2)}$, and using the corresponding off-sets. So in both cases (separate or common ϵ), the update step can be carried out with a simple univariate logistic regression maximum likelihood estimator. Computing a common ϵ in the case that we use linear regression for Y and logistic regression for $L(1)$ requires some programming.

We note that the clever covariate changes at each update step since the estimator of Q is updated at each step and the clever covariate is defined by the current Q-fit. Let $Q^*_{L(j)n}$, $j = 1, 2$, and Q^*_n denote the final update (at convergence of the MLE of ϵ to zero) of $Q_{L(j)n}$, $j = 1, 2$, and Q_n, respectively. The T-MLE of ψ is now given by $\Psi(Q^*_n)$.

A one-step T-MLE: Interestingly, if we use a separate $\epsilon_{L(j)}$ for $j = 1, 2$, first carry out the tmle update for $Q_{L(2)n}$, and use this updated $Q^*_{L(2)n}$ in the targeted MLE update for $Q_{L(1)n}$, then we obtain a targeted MLE-algorithm that converges in two simple steps, representing a single step update of Q_n. Below, we will generalize this one-step targeted MLE algorithm for updating an initial Q_n for general longitudinal data structures.

Statistical inference for T-MLE: Let $D^*(Q, g)$ be the efficient influence curve at $p_{Q,g} = Q^*g$, as defined in the above Theorem. Under regularity conditions, the T-MLE is consistent and asymptotically linear with influence curve $D^*(Q^*, g_0)$, where Q^* denotes the limit of Q^*_n, and g_0 is the true treatment mechanism. As a consequence, for construction of confidence intervals and testing one can use as working model $\psi^*_n \sim N(\psi_0, \Sigma_0)$, where $\Sigma_0 = ED^*(Q, g_0)^2$ is the variance of the efficient influence curve at (Q^*, g_0). Here Σ_0 can be estimated with the empirical covariance matrix of $D^*(Q^*_n, g_0)(O_i)$, $i = 1, \ldots, n$.

Targeted Loss-based selection among T-MLE's indexed by different initial estimators: Sequentially randomized trials allow us to select a targeted loss function for selection among different targeted maximum likelihood estimators indexed by different initial estimators. For the sake of illustration, we assume ψ_0 is one-dimensional. Suppose a collection of initial estimators is available for Q_0. Let $Q^*_{kn} = Q^*_k(P_n)$ be the corresponding targeted maximum likelihood estimators, $k = 1, \ldots, K$. One of these initial estimators might correspond with a super learner based on the log-likelihood loss function. We can select among these targeted maximum likelihood esti-

mators based on cross-validated risk of the loss function

$$L(Q) \equiv D^*(Q, g_0)^2,$$

which is indeed a valid loss function since it satisfies $Q_0 = \arg\min_Q E_0 L(Q)(O)$ among all Q with $\Psi(Q) = \psi_0$. The latter loss function is now a loss function for the whole Q and is very targeted towards ψ_0 since it corresponds exactly with the asymptotic variance of the targeted MLE. Thus, we would select k with the cross-validation selector:

$$k_n = \hat{k}(P_n) = \arg\min_k E_{B_n} P^1_{n,B_n} D^*(Q^*_k(P^0_{n,B_n}), g_0)^2,$$

where $B_n \in \{0,1\}^n$ denotes a random vector of binaries indicating a split in training sample $\{i : B_n(i) = 0\}$ and validation sample $\{i : B_n(i) = 1\}$, P^0_{n,B_n}, P^1_{n,B_n}, are the corresponding empirical distributions of the training and validation sample. Here we used the notation $Pf \equiv \int f(o)dP(o)$. The selected targeted maximum likelihood estimator is then $Q^*_n \equiv Q^*_{k_n}(P_n)$, and ψ_0 is now estimated with the substitution estimator $\Psi(Q^*_n)$.

Assuming a uniformly bounded loss function (i.e., a uniform bound on the efficient influence curve), due to oracle results of the cross-validation selector, the resulting targeted maximum likelihood estimator $\Psi(Q^*_n)$ will be at least as efficient as any of the candidate targeted maximum likelihood estimators $\Psi(Q^*_{kn})$, $k = 1, \ldots, K$.

Construction of Targeted Initial estimators: Above we showed that the projection of the efficient influence curve on the tangent space of the conditional distribution of $L(1)$ can be written as $C_{L(1)}(L(1) - Q_{L(1)})$, and for $Y = L(2)$, as $C_{L(2)}(L(2) - Q_{L(2)})$, where we use short-hand notation. For the purpose of constructing an initial estimator of $Q_{L(1)}$, we can use loss-based learning based on the weighted squared-error loss function $L_1(Q_{L(1)}) = C^2_{L(1)}(L(1) - Q_{L(1)})^2$, and, similarly, for the purpose of constructing of an initial estimator of $Q_{L(2)}$, we can use loss-based learning based on the weighted squared-error loss function $L_2(Q_{L(2)}) = C^2_{L(2)}(Y - Q_{L(2)})^2$. These are targeted loss functions since they correspond with the components of the variance of the efficient influence curve. Since the clever covariate $C_{L(2)}$ only depends on g_0, the required weights $C^2_{L(2)}$ for loss-based learning of $Q_{L(2)}$ are completely known. Therefore, we first apply the loss-based learning of the true $Q_{L(2)0}$. Let, $Q_{L(2)n}$ be the resulting estimator. Now, we plug such an estimator into the weight-function $C_{L(1)}$, and we use the resulting weights $C^2_{L(1)}$ to apply loss-based learning of $Q_{L(1)}$. In this way, using this

backwards sequential loss-based learning, we can generate initial candidate estimators of $Q_{L(1)}, Q_{L(2)}$ that are themselves already targeted by being based on these weighted squared-error loss functions (e.g. using different regression algorithms but using the weight option). We can now select among the targeted MLE indexed by these different targeted initial estimators, by using cross-validation with the above mentioned loss function $L(Q) = D^*(Q, g_0)^2$.

As might already be apparent, and certainly becomes apparent in the next section, this powerful approach combining loss-based learning with targeted MLE for the analysis of the simple two-stage sequentially randomized controlled trial generalizes to all sequentially randomized controlled trials for any target parameter, any number of stages, and higher dimensional intermediate time-dependent covariates.

We remark that the above targeted maximum likelihood estimator can also be applied to the data structure $L(0), A(0), L(1), \Delta, L(2) = \Delta Y$, where $A(0)$ is a treatment assigned at baseline (e.g, RCT), $L(1)$ represents the data collected between baseline and the time point at which the outcome Y is measured, and Δ is a missing indicator for Y. One simply applies the above data structure with $A(2) = \Delta$. Off course, if $L(1)$ is not binary, then the above estimator needs to be generalized as carried out in the next section, and, the missingness mechanism might need to be estimated from the data.

3 Targeted MLE of parameters of the G-computation formula.

We will now present the general approach to obtain a targeted maximum likelihood estimator, including the selection among different targeted maximum likelihood estimators indexed by different initial estimators. The choice of loss function we will use for the latter will depend on if one is willing to assume that the treatment/censoring mechanism is correctly estimated (or known, as in a S-RCT), or that one wishes to rely on double robustness, and we will provide appropriate loss functions for both purposes. This will generalize the above targeted maximum likelihood estimator for a two-stage sequentially randomized controlled trial to arbitrary sequentially randomized controlled trials, including S-RCT's that are subject to right-censoring or missingness and for which one is willing to assume that censoring/missingness is well understood. In addition, it will present the double robust T-MLE for

observational studies.

Organization: Firstly, we will present the likelihood using binary coding of the data structure O. Second, we will present a representation of the efficient influence curve based on this binary factorization of the likelihood. Third, we present the fluctuation/least favorable model of the initial estimate and the corresponding targeted maximum likelihood estimator. Fourth, we present a closed form one-step version of this targeted maximum likelihood estimator that applies if one is willing to fit a separate fluctuation parameter for each factor of the G-computation formula factor of the likelihood. Fifth, we present a targeted loss function that can be used to select among different targeted maximum likelihood estimators indexed by different initial estimators. We also present a particular type of targeted maximum likelihood estimator that uses a degenerate initial estimator for the intermediate factors of the G-computation formula, so that the targeted MLE algorithm only requires updating the final outcome conditional distribution. Finally, we make some observations regarding the pursuit of targeted dimension reductions simplifying the G-computation formula, which can form an important ingredient for generating different candidate targeted MLE's, and control complexity.

3.1 A factorization of likelihood of data in terms of binary variables.

Suppose the data structure for one experimental unit $O = (L(0), A(0), \ldots, L(K), A(K), L(K+1))$ involves collection of treatment and censoring actions coded with $A(t)$ at times $t = 0, \ldots, K$, and time-dependent covariate and outcome data at times $t = 0, \ldots, K+1$. We note that $L(t)$ can become degenerate after censoring and or after a terminal event like death, so that this data structure O also allows for longitudinal data structures that are truncated by the minimum of right-censoring and death. By choosing a fine enough discretization in time this data structure also approximates treatment and censoring processes $A(t)$ that evolve in continuous time.

For the sake of presentation, we will assume that $A(t)$ and $L(t)$ are discrete valued for all t so that the likelihood of O can be expressed in terms of probabilities, thereby avoiding technical difficulties regarding choice of dominating measure, without affecting the realm of practical applications.

The time ordering implies a graph with observed nodes $L(t), t = 0, \ldots, K+$

1, and $A(t)$, $t = 0, \ldots, K$, and a corresponding factorization of the observed data likelihood of O, given by

$$p_0 = \prod_{t=0}^{K+1} Q_{L(t)} \prod_{t=0}^{K} g_{A(t)},$$

where $Q_{L(t)}$ and $g_{A(t)}$ denote the conditional probability distributions of $L(t)$, given parents $Pa(L(t))$, and $A(t)$, given parents $Pa(A(t))$, respectively. The parent sets could be known to be subsets of the parent set implied by the time ordering of data structure, as discussed in introduction.

This factorized likelihood can be subjected to static and dynamic interventions on the $A()$ process mapping the probability distribution of O into probability distributions of O_d corresponding with a static or dynamic intervention d, often referred to as the G-computation formula. These interventions could involve all A-nodes as well as a subset of these nodes. The corresponding probability distributions of O_d are obtained by removing the $g_{A(t)}$'s corresponding with the $A(t)$ nodes on which an intervention is carried out under rule d, and substituting for $A(t)$ in the conditioning events (i.e., parents) of the $Q_{L(l)}$-factors with $l > t$ the corresponding intervened values.

In many applications $A(t) = (A_1(t), A_2(t))$ involves two types of actions $A_1(t)$ and $A_2(t)$, both relevant for defining the parameter of interest of the probability distribution of O. For example, $A_1(t)$ might be the treatment assigned at time t, $A_2(t)$ might be an indicator of being right-censored at time t, and the scientific parameter of interest, $\Psi(P_0)$, might be defined as a parameter of the distribution of O under the intervention on A defined by no-censoring at any time point, and a certain treatment intervention. In many cases, one defines the scientific parameter of interest in terms of changes of the latter distribution under different treatment regimens, and always no censoring: for example, marginal structural models for static or realistic dynamic treatment regimens provide such parameters, as we demonstrate in Section 5.

We will consider the case that for each node, the model for the conditional distributions of a nodes, given the parents, is nonparametric. Let Ψ be the parameter mapping so that $\psi_0 = \Psi(P_0)$ denotes the parameter of interest.

Without loss of generality, we assume that, for each $t \in \{1 \ldots, K+1\}$, $L(t)$ can be coded in terms of $n(t)$ binary variables $\{L(t, j) : j = 1, \ldots, n_t\}$,

so that $Q_{L(t)}$ can be further factorized as

$$Q_{L(t)} = \prod_{j=1}^{n(t)} Q_{L(t,j)},$$

where we define $Q_{L(t,j)}$ as the conditional distribution of $L(t,j)$, given its parents $Pa(L(t,j))$ defined as the parents of $L(t)$ augmented with the first $j-1$ variables $L(t,1),\ldots,L(t,j-1)$, $l=1,\ldots,j-1$. Note that this factorizations depends on a user-supplied ordering of the binary variables. For example, this particular coding and ordering might be implied by what is considered natural. The choice of coding and ordering does not affect the theoretical properties of the resulting targeted MLE, but it does imply the binary predictors $Q_{L(t,j)}$ one will need to estimate from the data.

This now provides the following likelihood factorization for the probability distribution of O:

$$p_0 = Q_{L(0)} \prod_{t=1}^{K+1} \prod_{j=1}^{n(t)} Q_{L(t,j)} \prod_{t=0}^{K} g_{A(t)}. \tag{1}$$

where $Q_{L(0)}$ denotes the marginal distribution of the baseline covariates $L(0)$.

3.2 General representation of efficient influence curve of target parameter.

We will now work out a general representation of the efficient influence curve we can apply to implement the targeted maximum likelihood estimator for general longitudinal data structures. These results provide us with a template for implementing the targeted maximum likelihood estimator for nonparametric models and essentially any type of longitudinal data structure that includes time dependent treatments and censoring actions that are realized in response to previously collected data.

Recall that in our model for P_0, for each node in the statistical graph, the conditional distribution is unspecified. Let Ψ be the parameter mapping so that $\psi_0 = \Psi(P_0)$ denotes the parameter of interest. If $\Psi(P_0) = \Psi^F(Q_0)$ is only a parameter of the Q_0, then we can present the efficient influence curve of Ψ as the projection of any influence curve (i.e., gradient of pathwise derivative) in the model in which g is known onto the tangent space of

Q (van der Laan and Robins (2003)):

$$D^* = \Pi(D \mid T_Q) \text{ for a certain gradient } D.$$

Such an estimating function D is often called an IPCW-estimating function (van der Laan and Robins (2003)). We will now be concerned with finding a representation of this efficient influence curve in terms of an orthogonal sum of scores of certain fluctuations $Q(\epsilon)$ of Q at $\epsilon = 0$, thereby implying a corresponding implementation of the targeted MLE.

The factorization (1) of the distribution P_0 implies an orthogonal decomposition of the tangent space at P_0 in our model, where this tangent space is a subspace of the Hilbert space $L_0^2(P_0)$ endowed with inner product $\langle h_1, h_2 \rangle = E_{P_0} h_1(O) h_2(O)$. This orthogonal decomposition of the tangent space $T(P_0) \subset L_0^2(P_0$ is given by

$$T(P_0) = T_{L(0)} + \sum_{t=1}^{K+1} \sum_{j=1}^{n(t)} T_{L(t,j)} + T_{CAR},$$

where $T_{L(0)}$ is the tangent space of $Q_{L(0)}$ consisting of the functions of $L(0)$ with mean zero, $T_{L(t,j)}$ is the tangent space of the conditional probability distribution $Q_{L(t,j)}$,

$$\begin{aligned} T_{L(t,j)} &= \left\{ V(L(t,j) \mid Pa(L(t,j))) - E_{Q_{L(t,j)}} V : V \right\} \\ &= \left\{ \{ V(1 \mid Pa(L(t,j))) - V(0 \mid Pa(L(t,j))\} \left(L(t,j) - Q_{L(t,j)}(1) \right) : V \right\}, \end{aligned}$$

and T_{CAR} is the tangent space of g. T_{CAR} can also be orthogonally decomposed as $\sum_{t=0}^{K} T_{A(t)}$ with $T_{A(t)}$ the tangent space of $g_{A(t)}$. Here we used the notation $E_{Q_{L(t,j)}} V = E(V \mid Pa(L(t,j)))$ for the conditional expectation w.r.t. $Q_{L(t,j)}$. If the parent sets are all implied by a specified ordering of all measured variables, then the model for P_0 is actually the nonparametric model so that the tangent space is saturated: $T(P_0) = L_0^2(P_0)$.

In the case that $\Psi(P_0)$ is a parameter of both Q_0 and g_0, the efficient influence curve D^* will also have components in T_{CAR}. An example of such a target parameter is $E(Y(1) - Y(0) \mid A = 1)$, the effect among the treated, based on observed data (W, A, Y) and the causal graph implied by time ordering W, A, Y. In that case the targeted MLE will also need to fluctuate an initial fit of g_0 with a fluctuation having a score that coincides with the efficient influence curve. For that purpose, let's also code $A(t)$ in terms of

binary variables. Let $A(t)$ be coded in terms of binary variables $\{A(t,j) : j = 1, \ldots, m(t)\}$, and consider the factorization

$$g_{A(t)} = \prod_{j=1}^{m(t)} g_{A(t,j)},$$

where an ordering needs to be specified so that the parents of $A(t,j)$ are given by the parents of $A(t)$ augmented with $A(t,1), \ldots, A(t,j-1)$.

The corresponding orthogonal decomposition of the tangent space of g is given by

$$T_{CAR} = \sum_{t=0}^{K} \sum_{j=1}^{m(t)} T_{A(t,j)}$$

where

$$\begin{aligned} &T_{A(t,j)} = \{V(A(t,j) \mid Pa(A(t,j))) - E(V \mid Pa(A(t,j))) : V\} \\ &= \left\{\{V(1 \mid Pa(A(t,j))) - V(0 \mid Pa(A(t,j))\}(A(t,j) - g_{A(t,j)}(1 \mid Pa(A(t,j))) : V\right\}. \end{aligned}$$

This factorization $p(O) = \prod_t \prod_j Q_{L(t,j)} \prod_t \prod_j g_{A(t,j)}$ yields the orthogonal decomposition of the tangent space $T(P_0)$ given by

$$T(P_0) = T_{L(0)} + \sum_{t=1}^{K+1} \sum_{j=1}^{n(t)} T_{L(t,j)} + \sum_{t=0}^{K} \sum_{j=1}^{m(t)} T_{A(t,j)}.$$

We can now state the corresponding Theorem for both a representation of a given efficient influence curve D^* as well as a projection of a function D, (e.g.) representing an inefficient influence curve for a parameter $\Psi(P) = \Psi^F(Q)$ in a model with g known, onto the tangent space T_Q of Q.

Theorem 2 *Consider the Hilbert space $L_0^2(P_0)$ and the factorization (1) of P_0. A function $D \in L_0^2(P_0)$ which is also an element of the tangent space $T(P_0)$ can be represented as*

$$D = D_{L(0)} + \sum_{t=1}^{K+1} \sum_{j=1}^{n(t)} D_{L(t,j)} + \sum_{t=0}^{K} \sum_{j=1}^{m(t)} D_{A(t,j)},$$

where

$$\begin{aligned} D_{L(0)} &= E(D \mid L(0)) - ED \\ D_{L(t,j)} &= \Pi(D \mid T_{L(t,j)}) \\ &= C_{L(t,j)}\{L(t,j) - Q_{L(t,j)}(1)\}, \end{aligned}$$

where

$$C_{L(t,j)} = E(D \mid L(t,j) = 1, Pa(L(t,j))) - E(D \mid L(t,j) = 0, Pa(L(t,j))),$$

for $t = 1, \ldots, K+1$, *and, for each* t, $j = 1, \ldots, n(t)$. *In addition,*

$$\begin{aligned} D_{A(t,j)} &= \Pi(D \mid T_{A(t,j)}) \\ &= C_{A(t,j)}\{A(t,j) - g_{A(t,j)}(1)\}, \end{aligned}$$

where

$$C_{A(t,j)} = E(D \mid A(t,j) = 1, Pa(A(t,j))) - E(D \mid A(t,j) = 0, Pa(A(t,j))).$$

In particular, the projection of D *onto the tangent space* T_Q *of* Q *can be represented as*

$$\Pi(D \mid T_Q) = D_0 + \sum_{t=1}^{K+1} \sum_{j=1}^{n(t)} D_{L(t,j)}.$$

If we represent D *as* $D(O) = D_1(A, L(A))/g(\bar{A} \mid X)$ *for some* D_1, $X = (L(a) : a)$, $L_a(t) = L_{\bar{a}(t-1)}(t)$, *assume that* $Pa(L(t,j))$ *includes* $\bar{A}(t-1)$, *then the above representation of* $\Pi(D \mid T_Q)$ *applies with*

$$\begin{aligned} &C_{L(t,j)} = \tfrac{1}{g(\bar{A}(t-1) \mid Pa(L(t)))} \times \\ &\left\{ E_Q\left(\textstyle\sum_{\bar{a}(t,K)} D_1 \mid L_a(t,j) = 1, Pa_a(L(t,j))\right) - E_Q\left(\textstyle\sum_{\bar{a}(t,K)} D_1 \mid L_a(t,j) = 0, Pa_a(L(t,j))\right)\right\}, \end{aligned}$$

Above, we used short-hand notation for $\sum_{\bar{a}(t,K)} D_1(\bar{A}(t-1), \bar{a}(t,K), L_{\bar{A}(t-1),\bar{a}(t,K)})$, *and* $\bar{a}(t,K) = (a(t), \ldots, a(K))$. *Here* $g(\bar{a}(t-1) \mid Pa(L(t)))$ *denotes the conditional probability of* $\bar{A}(t-1) = \bar{a}(t-1)$, *given* $Pa_{\bar{a}(t-1)}(L(t))$, *and it also equals the conditional probability of* $\bar{A}(t-1) = \bar{a}(t-1)$, *given* $L_a(t,j), Pa_a(L(t,j))$.

Proof of Theorem. We only need to prove the latter representation of $C_{L(t,j)}$.

We have $D(A, L(A)) = D_1(A, L(A))/g(\bar{A} \mid X)$, and we consider the case that $Pa(L(t,j))$ includes $\bar{A}(t-1)$. For the sake of this proof we exclude the treatment nodes $\bar{A}(t-1)$ from $Pa(L(t,j))$. Setting $\bar{A}(t-1) = \bar{a}(t-1)$, gives us the following conditional expectation to consider

$$E(D_1(\bar{A})/g(\bar{A} \mid X) \mid L(t,j), Pa(L(t,j)), \bar{A}(t-1) = \bar{a}(t-1)).$$

We first condition on X and $\bar{A}(t-1)$. This corresponds with taking an expectation w.r.t. $\prod_{s=t}^{K} g(A(s) \mid Pa(A(s)))$. This gives us

$$E(\sum_{\bar{a}(t,K)} D_1(\bar{a})/g(\bar{a}(t-1) \mid X) \mid L_{\bar{a}}(t,j), Pa_{\bar{a}}(L(t,j)), \bar{A}(t-1) = \bar{a}(t-1)).$$

This conditional expectation for each $\bar{a}(t,K)$-specific term is a sum over L_a compatible with $L_a(t,j), Pa_a(L(t,j))$. Specifically,

$$\begin{array}{l}
\sum_{L_a} \frac{D_1(\bar{a})}{g(\bar{a}(t-1)|X)} P(L_a \mid L_a(t,j), Pa_a(L(t,j)), \bar{A}(t-1) = \bar{a}(t-1)) \\
= \sum_{L_a} \frac{D_1(\bar{a})}{g(\bar{a}(t-1)|X)} \frac{P(L_a, L_a(t,j), Pa_a(L(t,j)), \bar{A}(t-1)=\bar{a}(t-1))}{P(L_a(t,j), Pa_a(L(t,j)), \bar{A}(t-1)=\bar{a}(t-1))} \\
= \sum_{\{L_a: L_a(t,j), Pa_a(L(t,j))\}} D_1(\bar{a}) \frac{P(L_a)}{P(L_a(t,j), Pa_a(L(t,j)), \bar{A}(t-1)=\bar{a}(t-1))} \\
= \frac{1}{g(\bar{a}(t-1)|L_a(t,j), Pa_a(L(t,j))} \sum_{\{L_a: L_a(t,j), Pa_a(L(t,j))\}} D_1(\bar{a}) \frac{P(L_a)}{P(L_a(t,j), Pa_a(L(t,j)))} \\
= \frac{1}{g(\bar{a}(t-1)|L_a(t,j), Pa_a(L(t,j))} E_Q(D_1(\bar{a}) \mid L_a(t,j), Pa_a(L(t,j)).
\end{array}$$

We will now prove that, by conditional independence assumptions of the statistical graph, $g(\bar{a}(t-1) \mid L_a(t,j), Pa_a(L(t,j))) = g(\bar{a}(t-1) \mid Pa_a(L(t)))$. To see this we first note that $g(\bar{a}(t-1) \mid L_a(t,j), Pa_a(L(t,j)))$ equals

$$\sum_{\bar{L}_a(t-1)} g(\bar{a}(t-1) \mid \bar{L}_a(t-1)) P(\bar{L}_a(t-1) \mid L_a(t,j), Pa_a(L(t,j))).$$

Since $Pa_a(L(t,j))$ are the parents of $L_a(t,j)$, we have $P(\bar{L}_a(t-1) \mid L_a(t,j), Pa_a(L(t,j))) = P(\bar{L}_a(t-1) \mid Pa_a(L(t,j)))$. Thus, this proves

$$g(\bar{a}(t-1) \mid L_a(t,j), Pa_a(L(t,j))) = g(\bar{a}(t-1) \mid Pa_a(L(t,j))).$$

More general, recall $L_a(t,j), Pa_a(L(t,j)) = L_a(t,1), \ldots, L_a(t,j), Pa_a(L(t))$, and note that $Pa_a(L(t))$ is included in $\bar{L}_a(t-1)$ (recall, that we excluded $\bar{A}(t-1)$ from $Pa(L(t,j))$ in this proof). We have $L_a(t,1), \ldots, L_a(t,j)$ is independent of $\bar{L}_a(t-1)$, given $Pa_a(L(t))$. So we obtain

$$\begin{array}{rcl}
P(\bar{L}_a(t-1) \mid L_a(t,j), Pa_a(L(t,j))) & = & P(\bar{L}_a(t-1) \mid L_a(t,1), \ldots, L_a(t,j), Pa_a(L(t))) \\
& = & P(\bar{L}_a(t-1) \mid Pa_a(L(t))).
\end{array}$$

This shows

$$g(\bar{a}(t-1) \mid L_a(t,j), Pa_a(L(t,j))) = g(\bar{a}(t-1) \mid Pa_a(L(t))).$$

To conclude, we have shown

$$\begin{array}{l}
E(\sum_{\bar{a}(t,K)} D_1(\bar{a})/g(\bar{a}(t-1) \mid X) \mid L_{\bar{a}}(t,j) = 1, Pa_{\bar{a}}(L(t,j)), \bar{A}(t-1) = \bar{a}(t-1)) \\
= \frac{1}{g(\bar{a}(t-1)|Pa_a(L(t)))} E(\sum_{\bar{a}(t,K)} D_1(\bar{a}) \mid L_{\bar{a}(t-1)}(t,j) = 1, Pa_{\bar{a}(t-1)}(L(t,j))),
\end{array}$$

which completes the proof. □

Remark about double robustness of efficient influence curve for general statistical graph: The efficient influence curve D^* at P depends on the Q-factor as well as a g representing conditional distributions of $A(t)$ nodes, possibly conditioning on subsets of the actual parents of $A(t)$. It is immediate that $P_0D^*(Q_0, g) = 0$ at possibly miss-specified g. To understand the possible additional robustness $P_0D^*(Q, g_0)$ for Q with $\Psi(Q) = \Psi(Q_0)$ and correctly specified g_0, and thereby the so called double robustness of the efficient influence curve (van der Laan and Robins (2003)), we make the following observation. By our latter representation in the above theorem, we have

$$\begin{aligned} &D^*(Q,g) = \textstyle\sum_t 1/g(\bar{A}(t-1) \mid (Pa_a(L(t)) : a)) \\ &\Big\{E_Q(\textstyle\sum_{\bar{a}(t,K)} D_1(\bar{a}(t-1), \bar{a}(t,K)) \mid L_a(t) = L(t), Pa_{\bar{a}(t-1)}(L(t)) = Pa(L(t))) \\ &- E_Q(\textstyle\sum_{\bar{a}(t,K)} D_1(\bar{a}(t-1), \bar{a}(t,K)) \mid Pa_{\bar{a}(t-1)}(L(t)))\Big\}\Big|_{\bar{a}(t-1)=\bar{A}(t-1)}, \end{aligned}$$

where we also have that $g(\bar{A}(t-1) \mid (Pa_a(L(t)) : a)) = g(\bar{A}(t-1) \mid (L_a(t), Pa_a(L(t)) : a))$, as we showed in the proof above. If we now take the conditional mean, given $(L_a(t), Pa_a(L(t)) : a)$, within the $\sum_t$-summation, then this corresponds with integration over $g_0(\bar{A}(t-1) \mid (Pa_a(L(t)) : a))$. Thus at a correctly specified g_0, we obtain that $P_0D^*(Q, g_0)$ equals

$$E_{Q_0} \sum_t \sum_{\bar{a}} \{E_Q(D_1(\bar{a}) \mid L_a(t), Pa_a(L(t))) - E_Q(D_1(\bar{a}) \mid Pa_a(L(t)))\},$$

thereby giving us an expression that does only depend on the Q_0-factor of the distribution of the data (thus nothing to do anymore with the conditional treatment probabilities). Some additional structure is now needed on the statistical graph to have that the latter equals zero at miss-specified Q. In particular, if $Pa(L(t)) = \bar{A}(t-1), \bar{L}(t-1)$ represents the history according to the time-ordered sequence representing the longitudinal data structure O, it follows, through cancelation of terms, that the latter equals $E_{Q_0} \sum_{\bar{a}} D_1(\bar{a})$, thereby giving the wished result (corresponding with the double robustness results in van der Laan and Robins (2003) for nonparametric full data models).

Using normal error regression to model and fluctuate conditional final outcome distribution. Consider the case that $\Psi(Q_0)$ only depends on the conditional distribution of a final outcome $Y = L(K+1)$, given

its parents $Pa(Y)$ through its conditional mean, and that the projection of the efficient influence curve (or any other gradient in model with g known) onto the tangent space of this conditional distribution Q_Y can be written as $C_Y(Y - E_Q(Y \mid Pa(Y))$ for some function C_Y of its parents $Pa(Y)$. Then it follows that there is no need to factorize the conditional distribution of Y in binary conditional distributions, but one could model the conditional distribution of Y with a normal error mean regression, and fluctuate the mean by adding the clever-covariate extension ϵC_Y. This was explicitly illustrated in Section 2 for the targeted MLE of EY_d.

3.3 The targeted MLE based on the binary representation of density

In this subsection we will define the targeted MLE based on the representation (1) of the density of O in terms of the binary predictors $Q_{L(t,j)}$, and, for the sake of presentation, we assume that our target parameter is only a parameter of Q_0. Consider an initial estimator $Q_{L(t,j)n}$ of each $Q_{L(t,j)}$, $t = 1, \ldots, K+1, j = 1, \ldots, n(t)$. We will estimate the first marginal probability distribution $Q_{L(0)}$ of $L(0)$ with the empirical distribution of $L_i(0)$, $i = 1, \ldots, n$. Let Q_n denote the combined set of $Q_{L(t,j)n}$ across all nodes $L(t,j)$.

The conditional distributions of $L(t,j)$ are binary distributions which we can estimate with machine learning algorithms (using logistic link) such as the super learner represented by a data adaptively (based on cross-validation) determined weighted combination of a user supplied library of machine learning algorithms estimating the particular conditional probability. These estimates could be obtained separately for each t, j or smoothing across time points t and or j could be employed if appropriate, by applying such machine learning algorithms to an appropriately constructed repeated measures data set. In particular, candidate estimators could be based on (guessed) subsets of $Pa(L(t,j))$.

In addition, let g_n be an estimator of g_0. We will now define the following fluctuations of the initial estimator $Q_{L(t,j)n} = Q_{L(t,j)}(P_n)$ of $Q_{L(t,j)}$:

$$\text{Logit}Q_{L(t,j)n}(\epsilon) = \text{Logit}Q_{L(t,j)n} + \epsilon C_{L(t,j)}(Q_n, g_n),$$

where we added the clever covariate $C_{L(t,j)n}$ obtained by substitution of the initial estimator Q_n and g_n of the true Q_0 and g_0.

One can now estimate ϵ with the MLE.

$$\epsilon_n = \arg\max_{\epsilon} \prod_{t=1}^{K+1} \prod_{j=1}^{n(t)} \prod_{i=1}^{n} Q_{L(t,j)n}(\epsilon)(O_i).$$

One could also obtain a separate MLE of ϵ for each factor indexed by (t, j):

$$\epsilon_{L(t,j)n} = \arg\max_{\epsilon} \prod_{i=1}^{n} Q_{L(t,j)n}(\epsilon).$$

One can now set $Q_n^1 = Q_n(\epsilon_n)$ to update Q_n. This updating process $Q_n^m = Q_n^{m-1}(\epsilon_n^m)$, $m = 1, \ldots,$ is now iterated till convergence, which defines the targeted MLE starting Q_n^* at initial estimator (Q_n, g_n).

We note that the $\epsilon_{L(t,j)n}$ for each factor separately can be estimated with standard logistic regression software using as off-set the logit of the initial estimator and having a single clever covariate $C_{L(t,j)}(Q_n, g_n)$. The single ϵ_n (uniform across t, j) defined above requires applying a single univariate logistic regression applied to repeated measures data set with one line of data for each factor indexed by (t, j), creating a clever covariate column that stacks $(C_{L(t,j)} : t, j)$ for each unit, and using the corresponding off-set covariate logit$Q_{L(t,j)n}$. So in both cases, the update step can be carried out with a simple univariate logistic regression maximum likelihood estimator using the off-set command (applied to a possibly repeated measures data set).

We note that the clever covariate changes at each update step since the estimator of Q is updated at each step and the clever covariate is defined by the current Q-fit in the iterative algorithm. Let $Q^*_{L(t,j)n}$ and Q_n^* denote the final update (at convergence of the MLE of ϵ to zero) of $Q_{L(t,j)n}$, and Q_n. The targeted MLE of ψ_0 is now given by $\Psi(Q_n^*)$.

3.4 A targeted MLE based on the binary predictor representation of density that converges in one step

In this section we will define a fast targeted MLE based on the representation (1) of the density of O in terms of the binary predictors $Q_{L(t,j)}$.

Consider an initial estimator $Q_{L(t,j)n}$ of each $Q_{L(t,j)}$, $t = 1, \ldots, K+1, j = 1, \ldots, n(t)$. We will estimate the first marginal probability distribution $Q_{L(0)}$

of $L(0)$ with the empirical distribution of $L_i(0)$, $i = 1, \ldots, n$. Let Q_n denote the combined set of $Q_{L(t,j)n}$ across t, j.

In addition, let g_n be an estimator of g_0. As above, we define the following fluctuations of the initial estimator $Q_{L(t,j)n}$ of $Q_{L(t,j)}$:

$$\text{Logit}Q_{L(t,j)n}(\epsilon) = \text{Logit}Q_{L(t,j)n} + \epsilon C_{L(t,j)}(Q_n, g_n),$$

where we added the clever covariate $C_{L(t,j)n}$ obtained by substitution of the initial estimator Q_n and g_n of the true Q_0 and g_0.

Monotone dependence on Q-property of the clever covariates: Consider the clever covariate representations of $C_{L(t,j)}$ presented in the above Theorem 2 for $Q_{L(t,j)}$ for the case that $D = D_1/g$ with D_1 not indexed by Q, g. Then the conditional expectations in the definition of the clever covariate $C_{L(t,j)}$ only depends on Q through $\{Q_{L(s,l)} : s > t, l\} \cup \{Q_{L(t,l)} : l > j\}$.

Let's enumerate all terms $Q_{L(t,j)}$ for $t \geq 1$ by moving row-wise: thus $Q_1 = Q_{11}$, $Q_2 = Q_{12}$, ..., $Q_{n(1)} = Q_{1n(1)}$, $Q_{n(1)+1} = Q_{21}$, and so on till $Q_N = Q_{K+1,n(K+1)}$, where $N = \sum_{t=1}^{K+1} n(t)$. Here we used temporarily the notation $Q_{12} = Q_{L(1,2)}$ and so on. Recall that $Q_{L(0)}$, the marginal distribution of $L(0)$, does not need to be fluctuated, and is thus not considered here: we will always estimate $Q_{L(0)}$ with the empirical distribution, so that no fluctuation is needed. Under this ordering, the k-th clever covariate C_k only depends on Q through $Q_{k+1}, \ldots, Q_N$, $k = 1, \ldots, N$. In particular, C_N does not depend on Q at all, while C_{N-1} depends on Q_N only, C_{N-1} depends on Q_{N-1}, Q_N, and so on. We refer to this property of the clever covariates as the monotone dependence (on Q) property, which will have an immediate implication for the corresponding iterative T-MLE algorithm.

We denote this monotonicity property with $C_k(Q) = C_k(Q_{k+1}, \ldots, Q_N)$, where we suppress the dependence on g since in the targeted MLE algorithm presented below g will not be updated.

We obtain a separate MLE of ϵ for each factor, but we start with last factor first, and use the update of last factor in clever covariate of $N-1$-th factor, carry out update of $N-1$-th factor, use the update of $N-1$-th factor in clever covariate of $N-2$-th factor, and so on till we update the first factor based on first clever covariate including all previously obtained updates. One could now start over, since Q_n got updated during this particular round of updating steps, and apply the same round of updating steps to the update of Q_n, and iterate this till convergence. The below Theorem states that this is not necessary, since the algorithm has converged after one round.

We state here the one step convergence of this targeted MLE algorithm.

Theorem 3 *Consider the targeted MLE algorithm above applied to an initial estimator Q_n, g_n, using a separate $\epsilon_{L(t,j)n}$ for each factor $Q_{L(t,j)n}$, $t \geq 1$, carrying out the updating steps one at the time, starting with final factor in likelihood, and going backwards till first term always incorporating the latest updates on Q_n, and Q_{0n} is the empirical distribution of $L_i(0)$, $i = 1, \ldots, n$. We can refer to one round of updating starting at final factor and ending at first factor as one step. This process can be iterated thereby defining an iterative algorithm.*

Suppose that for each $t \geq 1, j$ the clever covariate in this algorithm, $C_{L(t,j)}(Q)$, only depends on Q through $Q_{sl} = Q_{L(s,l)}$ for $s > t$ and for $s = t, l > j$. In that case, the above iterative targeted MLE algorithm converges in one step/round, and thus in exactly $N = \sum_{t=1}^{K+1} n(t)$ updating steps.

We recall from the previous Theorem, if $D(O) = D_1(O)/g(\bar{A} \mid X)$, and the probability distribution of O is factored in binary predictors as in (1), then $D^* = \Pi(D \mid T_Q) = D_0 + \sum_{tj} D_{tj}$, where $D_{tj} = C_{L(t,j)}(L(t,j) - Q_{L(t,j)}(1))$, and

$$C_{L(t,j)} = \frac{1}{g(\bar{A}(t-1)|X)} \times$$
$$\left\{ E_Q \left(\sum_{\bar{a}(t,K)} D_1 \mid L(t,j) = 1, Pa(L(t,j)) \right) - E_Q \left(\sum_{\bar{a}(t,K)} D_1 \mid L(t,j) = 0, Pa(L(t,j)) \right) \right\}.$$

Here we used short-hand notation for $\sum_{\bar{a}(t,K)} D_1(\bar{A}(t-1), \bar{a}(t,K), \bar{L}_{\bar{a}(t,K)}(K+1))$, and $\bar{a}(t,K) = (a(t), \ldots, a(K))$.

This monotonicity property of the clever covariate holds if D_1 does not depend on Q itself. More generally, it holds if

$$D(Q) = \frac{D_1 + C_1(Q)}{g}, \; C_1(Q) \text{ is only function of } O \text{ through } L(0), \bar{A}(K),$$

(so that $C_1(Q)$ will cancel out in the representation of $C_{L(t,j)}$) and D_1 does not depend on Q (it can depend on g).

This Theorem allows us to define closed form targeted MLE algorithms for a large class of parameters in our semiparametric model defined by no constraints on any of the conditional node specific distribution, given their specified parent nodes. To utilize this closed form one-step targeted MLE, one is forced to carry out a separate update step for each factor (only once), but one can still use smoothing across many factors for the initial estimator.

3.5 Targeted loss-based selection among targeted MLE.

The basic idea is as follows. All our candidate estimators of Q_0 are targeted maximum likelihood estimators, indexed by different initial estimators of Q_0, and using same g_n of g_0. Due to fact that these targeted MLE's solve the efficient influence curve equation, it follows that the bias for ψ_0 involves a product of $Q_n^* - Q_0$ and $g_n - g_0$: see asymptotic linearity Theorems in van der Laan and Robins (2003) and van der Laan and Gruber (2009). The goal is clearly to estimate Q_0 as accurately as possible, which will maximize efficiency and minimize bias for ψ_0. Therefore, we want to use cross-validation to select among different targeted maximum likelihood estimators, using a loss function whose risk is minimized at Q_0. However, there are many choices for the loss-based dissimilarity, $E_0 L(Q) - L(Q_0)$, between a candidate Q and Q_0 possible, and one will be more targeted towards ψ_0 than another. For example, we can use the log-likelihood loss function, a penalized log-likelihood loss function presented in (van der Laan and Gruber (2009)), and other loss functions inspired by the efficient influence curve of ψ_0, as presented here (see also van der Laan and Gruber (2009)).

Here we present two loss functions for Q_0 that are identified by the efficient influence curve of Ψ. Firstly, if g_0 is known , then we can use

$$L_1(Q) = \{D^*(Q, g_0)\}^2.$$

If $D^*(Q, g_0) = D(Q, g_0) - \Psi(Q)$, then it follows that indeed $Q_0 = \arg\min_Q E_0 L_1(Q)$, since the variance under P_0 of $D^*(Q, g_0, \psi_0)$ is minimized at $Q = Q_0$ (van der Laan and Robins (2003)). For more general efficient influence curves, the latter property has to be explicitly verified: at minimal, if $D^*(Q, g) = D(Q, g, \Psi(Q))$, then one can replace $\Psi(Q)$ by a consistent estimator of ψ_0, and use the loss function $D^2(Q, g_0, \psi_n)$. By the argument above, the loss function is still valid if one is willing to assume a consistent and good estimator of g_0 (an estimator that will converge faster to true g_0 than the estimators of Q_0 will converge to Q_0).

To explain the rationale of this loss function, first consider the case that g_0 is known. If g_0 is known, a targeted MLE for which Q_n^* converges to Q with $\Psi(Q) = \psi_0$ is asymptotically linear with influence curve $D^*(Q, g_0)$ (van der Laan and Rubin (2006)) and it is well known that the variance of $D^*(Q, g_0)$ for a Q with $\Psi(Q) = \psi_0$ is minimal at $Q = Q_0$, which then corresponds with the semiparametric information bound. Thus, $E_0 L_1(Q)$ equals the asymptotic variance of the influence curve of the targeted MLE. Under the assump-

tion that $L_1(Q)$ is uniformly bounded in all candidate Q's, we can apply the Theorems on the cross-validation selector (e.g. van der Laan and Dudoit (2003)), which proves that either the cross-validation selector is asymptotically equivalent with the oracle selector, or it achieves the parametric rate of convergence. As a consequence, loss-function based cross-validation based on this loss function will, for large enough sample size, select the targeted maximum likelihood estimator with the smallest asymptotic variance of its resulting substitution estimator of ψ_0 (excluding the case that there are ties). If g_0 is unknown, but estimated at a fast rate relative to the rate at which one estimates Q_0, then the above argument for the cross-validation selector still applies in first order: see van der Laan and Dudoit (2003) for oracle results for the cross-validation selector based on loss functions with nuisance parameters. If g_0 is estimated, and $Q \neq Q_0$, then the influence curve of the targeted MLE involves another contribution, reducing the variance relative to the variance of the influence curve for g_0 known. In this case, $L_1(Q)$ is not exactly the asymptotic variance of the targeted MLE, but it is still minimized at the optimal Q_0, and it represents a large component of the true asymptotic variance of the targeted MLE.

Consider now the case that we are *not* willing to assume that estimation of g_0 is easy relative to estimation of Q_0. In that case, the above loss function is not appropriate. Recall the representation of the efficient influence curve $D^* = D_0 + \sum_{t,j} D^2_{L(t,j)}$ with $D_{L(t,j)} = C_{L(t,j)}(L(t,j) - Q_{L(t,j)}(1))$. We make the following observation (using short-hand notation):

$$\begin{aligned} \mathrm{VAR}(D^*(Q_0, g_0)) &= ED^2_{L(0)} + \sum_{t,j} E_0 D^2_{L(t,j)} \\ &= ED^2_{L(0)} + \sum_{tj} E_0 C^2_{L(t,j)}(L(t,j) - Q_{L(t,j)}(1))^2, \end{aligned}$$

This suggests to use as loss function for $Q_{L(t,j)}$, $t \geq 1$, the weighted squared-error loss function:

$$L_2(Q) = \sum_{tj} C^2_{L(t,j)}(L(t,j) - Q_{L(t,j)}(1))^2,$$

which is indexed by a weight function $C^2_{L(t,j)}$. One would need to obtain an initial estimator of these weights which depend on both Q_0 and g_0. However, note that even if we estimate these weights inconsistently, then this loss function $L_2(Q)$ remains a valid loss function for Q_0, thereby preserving the

double robustness of the resulting targeted maximum likelihood estimator of Q_0.

In van der Laan and Gruber (2009) other loss functions implied by the efficient influence curve are proposed, including the variance of efficient influence curve at a *collaborative* estimator of g_0.

3.6 The targeted-MLE at a degenerate initial estimator for intermediate time-dependent covarariate factors.

Consider the likelihood factorization as used to define the G-computation formula, and assume that the IPCW estimating function is of the form stated in the above Theorem 3. If one of the node-specific conditional distribution is estimated with a *degenerate* conditional distribution, given the data generated by previous node-specific conditional distributions, then Theorem 3 implies that the projection of a function of O on the tangent space generated by that factor equals zero.

For example, suppose the likelihood is factored according to the ordering $L(0), A(0), L(1), A(1), Y$. The projection of a function $D(O)$ onto the tangent space of $Q_{L(1)}$ is zero at a degenerate $Q_{L(1)}$, even if the true conditional distribution of $L(1)$ is not degenerate.

This insight suggests a simple-to-compute version of targeted MLE. Suppose we obtain an initial estimator Q_n^0 that is degenerate for all factors except the last one, and we use the empirical distribution for the marginal distribution of the baseline covariates. In that case, only the last factor, say $Q_{Y=L(K+1)}$, needs to be updated in the targeted MLE algorithm. As a consequence, the targeted MLE requires only one update, and thus converges in one single updating step.

Thus, in this case one estimates most of the system with a deterministic system, and only the last factor is estimated with a non-degenerate conditional probability distribution that is updated with a clever covariate depending on the treatment mechanism. Due to the double robustness of the targeted MLE, the resulting targeted MLE will be consistent and asymptotically linear if the treatment mechanism is correctly specified, and will still gain efficiency if the degenerate distributions are doing a reasonable job: since the degenerate distribution will be misspecified it is not reasonable anymore to rely on correct specification of the initial estimator Q_n^0 of Q_0 for consis-

tency. Note also that this simplified targeted ML estimator still allows us to apply the collaborative targeted MLE approach for selection among different treatment mechanism estimators based on the log-likelihood of the targeted estimator Q_n^1 indexed by the treatment mechanism estimator: see van der Laan and Gruber (2009).

We can view this particular simple targeted maximum likelihood estimator as one particular candidate among a set of candidate targeted maximum likelihood estimators, and use loss-based cross-validation to select among these candidate targeted maximum likelihood estimators. One would use one of our proposed (efficient-influence curve based) loss functions, such as the weighted squared error loss function, since the log-likelihood loss function will become undefined as a degenerate distribution.

3.7 Dimension reduction for time-dependent covariates.

One could use a loss function on the Q-factor of the binary coded complete data structure, and use loss-function based cross-validation to select among different fits, thereby implicitly carrying out a dimension reduction. For example, by not including a node in the graph in parent sets of other nodes it will be equivalent to removing the node from the data structure, and such moves can be scored based on the loss function. In this manner one might build an initial estimator Q_n whose G-computation formula for parameter of interest is only affected by conditional distribution of subset of all nodes, thereby also simplifying the targeted MLE update.

Here we wish to investigate alternative targeted dimension reductions that would, in particular, reduce the computational complexity of the targeted maximum likelihood estimator which is driven by the number of binary variables coding the data structure. This reduced complexity/dimension can also imply that the loss function for the Q_0 of the reduced data structure implies a more targeted dissimilarity for the purpose of fitting $\Psi(Q_0)$.

If a multivariate $L(t)$ is reduced to a one dimensional time-dependent covariate, then the targeted maximum likelihood estimator is simpler, but if this reduction means that $A(t)$ now depends on measured variables beyond the one dimensional reduced time-dependent covariate, then this reduction will also have caused bias. In addition, much information might have been lost, thereby causing variance. So care is needed.

Let's revisit the two-stage sequentially randomized controlled trial with data structure $O = L(0), A(0), L(1), A(1), Y = L(2)$, but let's now consider the more general case that $L(1)$ can be a multivariate vector with discrete and/or continuous components. Suppose that we wish to estimate $EY(1,1)$. Inspection of the efficient influence curve of $EY(1,1)$ shows that it only depends on the conditional distribution of Y through its mean $E(Y \mid A(0) = 1, A(1) = 1, \bar{L}(1))$. This suggest that $L_Q(1) = E(Y \mid A(0), A(1) = 1, \bar{L}(1))$ denotes a targeted dimension reduction: below we provide a general approach which implies this precise dimension reduction. In addition, let $L_g(1)$ be defined as the propensity score $P(A(1) = 1 \mid L(0), A(0), L(1))$. A targeted dimension reduction of the observed O is now given by $(L(0), A(0), L_Q(1), L_g(1), A(1), Y)$. We can fit both $L_Q(1)$ and $L_g(1)$ from the data using super-learning, thereby obtaining an estimated dimension reduction O^r. A targeted MLE for this (estimated) reduced data structure now involves fitting $Q_{L_Q(1)}$, $Q_{L_g(1)}$, and Q_Y, where only the conditional mean of Y is needed. However, by definition of $L_Q(1)$, the conditional mean of Q_Y at $A(1) = 1$ equals $L_Q(1)$, suggesting that we can exclude $L_g(1)$ from the parent set of Y without meaningful loss of information. Then, the conditional distribution $Q_{L_g(1)}$ does not affect the G-computation formula of the distribution of $Y(1,1)$ or, more general, the joint distribution of $Y(1,1)$ and $L(0)$. As a consequence, in this case we do not even need to fit $Q_{L_g(1)}$.

To summarize, in this manner we have succeeded in dramatically reducing the complexity of a targeted MLE by replacing the fitting of a conditional distribution of a multivariate random variable $L(1)$ into fitting of a univariate conditional distribution of $L_Q(1)$.

Let's now generalize this type of targeted dimension reduction procedure. Consider a general longitudinal data structure $L(0), A(0), \ldots, L(K), A(K), Y = L(K+1)$, and let's consider the case that $A(j)$ is binary, $j = 0, \ldots, K$. The dimension reduction can be guided by the actual form of the efficient influence curve for the target parameter. To demonstrate this, we first note that the efficient influence curve can often be represented as $D_{IPCW}(g_0, \psi_0)(L(0), A, Y) - \sum_{j=0}^{K} E(D_{IPCW} \mid A(j), Pa(A(j))) - E(D_{IPCW} \mid Pa(A(j)))$ for some IPCW-estimating function (see van der Laan and Robins (2003)). The latter differences of two conditional expectations can also be written as $C(j)(A(j) - P(A(j) = 1 \mid Pa(A(j))$, where

$$C(j) = E(D_{IPCW} \mid A(j) = 1, Pa(A(j))) - E(D_{IPCW} \mid A(j) = 0, Pa(A(j))).$$

For example, if $\psi_0 = EY(1)$, then $D_{IPCW}(O) = \{I(\bar{A} = 1)/g(\bar{A} \mid X)\}Y - \psi_0$.

As we did before, we can factorize this difference of conditional expectations in terms of a factor only depending on Q_0 and a factor only depending on g_0. We can define $L_Q(j)$ as the Q_0-factor only, thereby preserving double robustness of the resulting targeted MLE. In addition, we define

$$L_g(j) = P(A(j) = 1 \mid Pa(A(j)).$$

If the target parameter is $EY(a)$ for a static regimen a, it follows that the efficient influence curve depends on O through the reduction

$$O^r = (L(0), A(0), L_Q(1), L_g(1), A(1), \ldots, L_Q(K), L_g(K), A(K), Y.$$

If the target parameter is $EY(d)$ for a dynamic treatment rule d, then, one also needs to include the time-dependent covariate the rule d uses to assign treatments. To summarize, inspection of the efficient influence curve of the target parameter defines a reduction O^r in terms of two time-dependent covariate processes, one representing the treatment asssignment probabilities as functions of the past, and one based on the Q_0-functions making up the efficient influence curve. These time-dependent covariates depend on unknown Q_0 and g_0. We will estimate these time-dependent covariates, by estimating the treatment mechanism, and the required $L_Q(j)$. We can now apply the targeted MLE to this reduced data structure.

As in our previous example, suppose that for each of the conditional distributions of Y and $L_Q(j)$, $j = 1, \ldots, K$, we do not include any of the $L_g(j)$ in the parent sets. We suggest that this comes at little cost in efficiency. Under this condition, the conditional distributions $Q_{L_g(j)}$ do not affect the G-computation formula of the distribution of $Y(d)$ or, more general, the joint distribution of $Y(d)$ and $L(0)$. As a consequence, in that case we do not even need to fit $Q_{L_g(j)}$, $j = 1, \ldots, K$. To summarize, in this manner we can dramatically reduce the complexity of a targeted MLE by replacing the fitting of a conditional distribution of a multivariate random variable $L(j)$ into only fitting the univariate conditional distributions of $L_Q(j)$ and possibly the conditional distribution of another time-dependent covariate that is used to define the target parameter (e.g., treatment rule based on time dependent biomarker). Note that we will still fit the treatment mechanisms of $A(j)$ conditional on its parents (under O^r) including $L_g(j)$, and thus just fit $P(A(j) = 1 \mid Pa(A(j)))$ with $L_g(j)$ itself.

This dimension reduction still allows for the construction of a collaborative estimator g_n of g_0, given an estimator Q_n of Q_0^r, representing the

conditional distributions of $L_Q(j), L_g(j)$ and Y. This just requires applying the C-T-MLE algorithm as presented in van der Laan and Gruber (2009) to the log-likelihood for Q_0^r, thereby scoring a fit of g_0 with the log-likelihood (or other loss function) of the targeted MLE of Q_0^r corresponding with the fluctuation function implied by the candidate g_0-fit.

By using as loss function the variance of the influence curve of the targeted MLE we can still select among different targeted maximum likelihood estimators indexed by different dimension reductions of the type presented above, assuming that each of them puts the maximal effort in obtaining an unbiased estimator.

4 A general template for targeted MLE of parameters of the G-computation formula

We present a road map for the computation of the targeted MLE and statistical inference.

Code data: Represent the data on one unit as a time-ordered data structure

$$O = (L(0), A(0), L(1), A(1), \ldots, L(K), A(K), L(K+1)).$$

It is assumed that $L(t)$ occurs before $A(t)$, and we are interested in effect of interventions on the A-nodes of this graph.

Define target parameter: Let P_0 be the probability distribution of O, and let $\psi_0 = \Psi(P_0)$ be the target parameter of interest. The probability distribution of O factorizes as $p_0 = Q_0 g_0$, where $Q_0 = \prod_{t=0}^{K+1} Q_{0L(t)}$ and $g_0 = \prod_{t=0}^{K} g_{0A(t)}$, $Q_{0L(t)}$ is the conditional distribution of $L(t)$, given $Pa(L(t)) = \bar{L}(t-1), \bar{A}(t-1)$, and $g_{0A(t)}$ is the conditional distribution of $A(t)$, given $Pa(A(t)) = \bar{L}(t), \bar{A}(t-1)$. If it is known that the parent sets of these nodes are smaller, then these smaller parent sets need to be enforced. We will assume that each of these conditional distributions is unspecified. Typically we will have that $\psi_0 = \Psi^F(Q_0)$ is only a parameter of the Q-factor of density p_0 of O. In causal inference most target parameters can be defined as a parameter of a distribution obtained by intervening on the A nodes in the complete system, which is thereby only a function of Q, i.e., the G-computation formula.

Determine efficient influence curve: In order to carry out the targeted MLE to estimate ψ_0 one will need to know the efficient influence curve $D^*(Q,g)$ for any (Q,g) identifying a distribution of the data. If ψ_0 only depends on P_0 through Q_0, then one can find an influence curve D_{IPCW} of Ψ in the model in which g_0 is known. Such influence curves can often be represented as so called inverse of probability of censoring weighted functions of a full data efficient influence curve (see van der Laan and Robins (2003) for a formal treatment of IPCW-estimating functions). In that case the efficient influence curve can be represented as a projection of such an IPCW-estimating function D_{IPCW} onto the tangent space of Q:

$$D^*(Q,g) = \Pi(D_{IPCW}(Q,g) \mid T_Q),$$

where T_Q is the tangent space of the Q-factor of the density p_0 of O. Our formulas of the clever covariates in the fluctuation function of the various factors in Q will be a direct function of this D_{IPCW}. Since for most target parameters the IPCW-estimating function is well known and easily constructed, this provides us with a straightforward way to obtain the right formulas for the clever covariates needed to define the fluctuation function of the targeted MLE step.

Determine binary factorization of likelihood: Consider a $L(t)$. Suppose $L(t) = (L(t,j) : j = 1,\ldots,n(t))$ consists of $n(t)$ components, which we denote with $L(t,j)$ for different j. Firstly, we determine a particular ordering, allowing us to model

$$Q_{L(t)} = \prod_{j=1}^{n(t)} Q_{L(t,j)},$$

where $Q_{L(t,j)}$ is the conditional probability distribution of $L(t,j)$, given $Pa(L(t,j)) = \bar{L}(t-1), L(t,1), \ldots, L(t,j-1), \bar{A}(t-1)$. It now remains to further factorize $Q_{L(t,j)}$. If $L(t,j)$ is binary, then we do not further factorize $Q_{L(t,j)}$. If $L(t,j)$ is a categorical variable with $n(t,j)$ categories, then assume an ordering of the categories $l = 1,\ldots,n(t,j)$, and factorize $Q_{L(t,j)}$ as

$$Q_{L(t,j)} = \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

where $Q_{L(t,j,l)}$ is the conditional distribution of the indicator $L(t,j,l)$ of $L(t,j) = l$, given $I(L(t,j) = m)$, $m = 1, \ldots, l-1$, and $Pa(L(t,j))$, where it is assumed that if one of these indicators $I(L(t,j) = m)$ for $m = 1, \ldots, l-1$ equals 1, then $Q_{L(t,j,l)}$ is degenerate at 0. Let $Pa(L(t,j,l))$ denote the parent set for this node $L(t,j,l)$.

If $L(t,j)$ is an ordered variable with $n(t,j)$ values, then we already have an ordering, and factorize $Q_{L(t,j)}$ as

$$Q_{L(t,j)} = \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

where $Q_{L(t,j,l)}$ is the conditional distribuiton of the indicator $L(t,j,l)$ of $L(t,j) = l$, given $I(L(t,j) = m)$, $m = 1, \ldots, l-1$, and $Pa(L(t,j))$, where it is assumed that if one of these indicators $I(L(t,j) = m)$ for $m = 1, \ldots, l-1$ equals 1, then $Q_{L(t,j,l)}$ is degenerate at 0.

Note that the latter $Q_{L(t,j,l)}$ (conditional on the previous $l-1$ indicators all being zero) is identified by a so called hazard probability $Q_{L(t,j,l)}(1 \mid Pa(L(t,j,l))$, i.e., a probability of having the random variable fall at level l, conditional on being larger or equal than level l, and $Pa(L(t,j,l))$.

To conclude, we have the following factorization for $Q_{L(t)}$,

$$Q_{L(t)} = \prod_{j=1}^{n(t)} \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

and thereby the factorization for the Q-factor of the density of O,

$$Q = \prod_{t} \prod_{j=1}^{n(t)} \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)}$$

in terms of conditional distributions of binary variables $L(t,j,l)$, conditional on parent nodes $Pa(L(t,j.l))$.

Compute formulas for clever covariates: For each binary variable $L(t,j,k)$, $Pa(L(t,j,k))$ denotes the conditioning random variable:

$$Pa(L(t,j,k)) = \bar{L}(t-1), \bar{A}(t-1), L(t,1), \ldots, L(t,j-1), L(t,j,1), \ldots, L(t,j,k-1),$$

and the following formulas have been provided for general parent sets as well. Suppose that $D = D_{IPCW}$ can be represented as

$$D(Q) = \frac{D_1 + C_1(Q)}{g},\ C_1(Q) \text{ is only function of } O \text{ through } L(0), \bar{A}(K),$$

and D_1 does not depend on Q (it can depend on g).

The efficient influence curve can now be represented as $D^* = \Pi(D \mid T_Q) = D_0 + \sum_{t \geq 1, j, k} D_{tjk}$, where $D_0 = E(D^* \mid L(0))$, and for $t \geq 1$, $D_{tjk} = C_{tjk} \left\{ L(t,j,k) - Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k))) \right\}$, with

$$\begin{aligned} &C_{tjk}(Q,g) = \tfrac{1}{g(\bar{A}(t-1)|X)} \times \\ &\{E_{tjk}(Q)(1, Pa(L(t,j,k))) - E_{tjk}(Q)(0, Pa(L(t,j,k)))\}, \end{aligned}$$

where we defined, for $\delta \in \{0,1\}$,

$$E_{tjk}(Q)(\delta, Pa(L(t,j,k))) = E_Q \left(\sum_{\bar{a}(t,K)} D_1 \mid L(t,j,k) = \delta, Pa(L(t,j,k)) \right).$$

Here we used short-hand notation for $\sum_{\bar{a}(t,K)} D_1(\bar{A}(t-1), \bar{a}(t,K), \bar{L}(K+1))$, and $\bar{a}(t,K) = (a(t), \ldots, a(K))$.

Given an estimator Q_n and g_n, $C_{tjk}(Q_n, g_n)$ denotes the clever covariate defining the fluctuation function of the estimator $Q_{L(t,j,k),n}$ obtained by adding $\epsilon C_{tjk}(Q_n, g_n)$ on the logit scale.

Define clusters of Q-probabilities that need to considered for pooling: We now need to define an initial estimator Q_n^0 of Q_0. One could estimate each $Q_{L(t,j,k)}$ separately with a machine learning algorithm for each t, j, k. Such an estimator can be considered as one particular candidate, but likelihood/loss function based cross-validation needs to be employed in order to evaluate this type of estimator relative to other estimators. Overall, smoothing in time t and/or category l is often very sensible and will improve the overall performance of the estimator: i.e. it will typically reduce the variance significantly at relatively minor loss in bias. Therefore, the user should define clusters of grid points in the (t,j,k)-grid for which the estimators of the corresponding $Q_{L(t,j,k)}$ need to be considered for being pooled: the particular machine learning algorithm applied to this cluster might still decide to not smooth in certain

time-points or levels, but it will be guided by cross-validated risk using a specified loss function. Let $\tau_1, \ldots, \tau_L$ be such clusters. For example $\tau_1 = \{(t, 1, k) : t = 1, \ldots, K, k = 1, \ldots, n(t, 1) - 1\}$ might indicate that component indicated by $j = 1$ needs to be considered for smoothing in both time t across all time points and in its level k. So each cluster typically represents a particular ordered variable (e.g., CD4 count) and it might state that the estimation procedure needs to respect the fact that the hazard of this variable at different levels k and different time points might need to be smoothed across time and level. For a categorical variable, smoothing in the categories will make no sense and would thus be made clear by the definition of the cluster for that categorical variable, but that cluster might still suggest smoothing over time t.

Apply loss-based super learner to repeated measures data set to estimate each cluster of Q-probabilities:

For each cluster τ_l, we create a pooled data set which has for each unit as many rows as there are grid points in the cluster which will be used to fit all the conditional probabilities $Q_{L(t,j,k)}$ with $(t, j, k) \in \tau_l$. So we create a data set which has as columns a time stamp t, a j-stamp, and a level stamp l, an outcome column for the binary $L(t, j, k)$, various covariates extracted from $Pa(L(t, j, k))$, across all points $(t, j, k) \in \tau_l$, each grid point representing a line of data, thereby generating repeated measures type data for each unit of observation. Regarding extraction of covariates from $Pa(L(t, j, k))$, one needs to make sure that these covariates have the same meaning across these different grid points. So this step involves defining a list of extractions the histories $Pa(L(t, j, k))$, such as the most recent in time measurements on the particular variable indicated by j. If this is not possible for (say early time points), then that is an indications that some of these (t, j, k) should not have been included in the cluster and might thus need to be fitted separately.

Given this definition of a repeated measures data set corresponding with cluster τ_l, we can now apply the super learner or any other machine learning algorithm to fit the regression of the binary $L(t, j, k)$ onto (t, j, k) and these covariate extractions from $Pa(L(t, j, k))$, across (t, j, k). This requires a choice of loss function $L(Q)(O)$. One possibility is the log-likelihood loss function $-\sum_{(t,j,k)\in\tau_l} \log Q_{t,j,k}(O)$. We can use a potentially more targeted loss function given by the repeated

measures squared error loss function $L(Q_{t,j,k} : (t,j,k) \in \tau_l)(O)$ defined as

$$\sum_{(t,j,k)\in\tau_l} w(t,j,k)R(t,j,k)\left\{L(t,j,k) - Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k)))\right\}^2,$$

for some weight function $w(t,j,k)$. Here $R(t,j,k)$ denotes an indicator of $L(t,j,k)$ being at "risk" of changing value. If $R(t,j,k) = 0$, then $Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k)))$ is either known to be zero or one, so that this loss function will only evaluate the conditional probability of $L(t,j,k)$, given its parents and given that it is at risk of changing.

As discussed previously, a particular weight function $w(t,j,k)$ that makes the risk of the loss function close to variance of efficient influence curve, and thereby targets the super learner fit towards the parameter of interest, is given by the square of the clever covariate:

$$w(t,j,k) = C(t,j,k)^2.$$

Since this weight depends on Q, g itself, this weighted super learner would require a two stage procedure, first an unweighted super learner (or other machine learning algorithm) to estimate Q_0, and a subsequent weighted super-learner using the first stage estimator to estimate the clever covariates, and thereby the weights.

Estimate treatment/censoring mechanism: The likelihood for g can also be factorized in binary conditional probability distributions, and will typically involve fewer binary conditional probability distributions. One can use log-likelihood based machine learning (e.g., super-learning) to estimate g_0. Regarding the choice of loss function for g_0, one needs to realize that the estimator of g_0 is only used to estimate the clever covariates, and that fact might guide the choice of loss function for g_0 so that the resulting estimator of g_0 is well suited for estimation of the clever covariates.

Targeted MLE algorithm at given initial and treatment/censoring mechanism estimator: Suppose that we are given an initial estimator Q_n of Q_0, described above, and an estimator g_n of g_0.

Define the fluctuation of the initial estimator Q_n:

$$\text{logit}Q_{L(t,j,k),n}(\epsilon) = \text{logit}Q_{L(t,j,k),n} + \epsilon C_{tjk}(Q_n, g_n).$$

We now have a variety of possibilities regarding estimation of ϵ depending on how much we want to smooth the estimator of ϵ across (t, j, k).

Firstly, we could create a single pooled repeated measures data set in which each unit contributes a line of data for each t, j, k. One now creates columns for the time stamp t, the variable indicator j, the category/level indicator k, the outcome $L(t, j, k)$, the offset $Q_{L(t,j,k),n}$, and the clever covariate C_{tjk}. One could now fit ϵ with logistic regression using the off-set command, regressing the binary indicator $L(t, j, k)$ onto the clever covariate C_{tjk}, thereby obtaining a single estimator ϵ_n for each t, j, k. One now updates the initial estimator $Q_n^1 = Q_n^0(\epsilon_n)$, and this process is iterated till convergence. In this way, the targeted bias reduction is established with minimal extra fitting, and therefore this might be the preferred method relative to the alternatives considered below.

Alternatively, one could create an extra column that labels the cluster τ_l, and, for each l, one runs the same iterative algorithm as above but now only applied to the repeated measures data set corresponding with the $(t, j, k) \in \tau_l$, i.e., for which this cluster label column equals l. In this case, one obtains a separate estimator ϵ_{nl} for each cluster $l = 1, \ldots, L$, and this ϵ_{nl} is used for each $(t, j, k) \in \tau_l$. Note that in this approach one uses the same pooling to fit ϵ as was used in the initial estimator.

Finally, consider the ordering based on $L(t_1, j_1, k_1) < L(t_2, j_2, k_2)$ if and only if either $t_1 < t_2$, or, if $t_1 = t_2$, $j_1 < j_2$, or, if $t_1 = t_2, j_1 = j_2$, $k_1 < k_2$. Using this ordering, define disjoint and complementary clusters of points that represent an interval in the ordering: thus the intervals cover all points in the (t, j, k)-grid.

So, using this ordering any (t, j, k) can now be denoted with an integer $V(t, j, k) \in \{1, \ldots, N\}$, and a cluster now has to be of the form $[a, b]$ representing all integers between a and b, including a, b. A typical cluster will now only run over the integers corresponding with the k values for a particular variable indicated by t, j. For each interval one now runs the same iterative algorithm as above but now only applied to the repeated measures data set corresponding with the (t, j, k) with $V(t, j, k)$ in the interval. This iterative targeted ML algorithm thus only updates the $Q_{L(t,j,k),n}$ for $V(t, j, k)$ in the interval. However, we run

these interval-specific targeted MLE algorithms *sequentially* starting with the last interval in the ordering. After having run the last interval algorithm thereby updating the $Q_{L(t,j,k),n}$ at the end of the ordering, we update these $Q_{L(t,j,k),n}$, and run the next interval (going backwards!) with the updated Q_n. In this second interval targeted ML algorithm one updates the $Q_{L(t,j,k),n}$ corresponding with the second interval while fixing the already obtained fits from the first interval. After having run this second interval algorithm and having updated the corresponding $Q_{L(t,j,k),n}$, one runs the targeted ML algorithm for the third interval (going backwards) updating the $Q_{L(t,j,k),n}$ corresponding with this third interval, while fixing the already obtained fits of the first and second interval. One iterates this updating process till one arrives at the first interval, at which time the algorithm is finished and the updated Q_n^* is complete.

This algorithm uses the fact that the clever covariates used in an interval only depend on the $Q_{L(t,j,k),n}$ with $V(t,j,k)$ in the interval and to the right of that interval. As a consequence, an update of an interval does not affect the targeted ML algorithm for the intervals to the right of that interval. Thus, by first updating the $Q_{L(t,j,k),n}$ at the end of the ordering and moving backwards in the ordering we can finish the algorithm in one round. In other words, we exploit the monotonicity property of the clever covariates as presented in Theorem 3 above, that allowed a closed form backwards targeted MLE algorithm (converging in one step) if one uses a separate ϵ for each factor $Q_{L(t,j,k),n}$.

Selection among different estimators of the treatment/censoring mechanism: The Collaborative Targeted MLE.

The targeted ML update of the initial super-learning fit Q_n is a function of the clever covariates and thereby depends on the choice of treatment/censoring mechanism estimator. Different choices of treatment/censoring mechanism estimator result in different clever covariates sets $(C_{tjk} : t, j, k)$ and thereby result in different increases in likelihood fits due to targeted MLE update.

As in van der Laan and Gruber (2009), we suggest to use log-likelihood-based cross-validation to select among a sequence of targeted maximum likelihood estimators using increasingly nonparametric estimators of the treatment mechanism to estimate the clever covariates, thereby

fine-tuning the depth of bias-reduction pursued. For details about such procedures we refer to van der Laan and Gruber (2009), including the fact that the resulting targeted maximum likelihood estimator is now collaborative double robust. Collaborative double robustness means that the targeted maximum likelihood estimator is consistent if one uses an estimator g_n that converges to a true censoring mechanism that correctly adjusts the covariates that explain (i.e. increase the likelihood fit relative to Q_n) the residual bias $Q - Q_0$, where Q denotes the limit of the initial estimator Q_n. That is, covariates that are not helpful in explaining residual bias $Q_n - Q_0$ do not need to be adjusted for in the censoring mechanism inputted in the clever covariate/least favorable model. One particular collaborative targeted MLE algorithm presented in van der Laan and Gruber (2009) corresponds with using a greedy forward selection building of the treatment/censoring mechanism (adding one covariate at the time) based on the penalized log-likelihood of the corresponding targeted MLE, and choosing the size of the g_0-fit (i.e. number of steps in forward selection algorithm) with penalized log-likelihood based cross-validation, using a penalty to stabilize the procedure in sparse data situations in which clever covariates can reach large/outlier values.

In order to save computer time one could decide to not cross-validate the initial estimator and also not cross-validate the treatment mechanism estimators in this cross-validated risk of this loss function. So the initial estimator is treated as a fixed off-set, and the clever covariates indexed by different treatment mechanism estimators are treated as given as well. The cross-validation thus concerns running the second stage targeted MLE algorithm described above on a training set, while fixing the treatment mechanism estimator (and thereby the clever covariate) and the offset-initial estimator at their fit based on the whole sample. In van der Laan and Gruber (2009) we point out that using a cross-validated initial estimator as off-set, in the sense that O_i is coupled with an initial estimator using the training sample excluding O_i, $i = 1, \ldots, n$, can be important to obtain the wished bias reduction with the targeted MLE in the case that the initial estimator is an overfit.

In addition, to save computer time, when carrying out this building and selection among different targeted maximum likelihood estimators we suggest that one might replace the candidate targeted maximum

likelihood estimators with one-step (or few steps) targeted maximum likelihood estimators only carrying out one ϵ_n-updating step. Once a targeted maximum likelihood estimator is selected, it is fully iterated till convergence, and the latter true targeted maximum likelihood estimator is the reported estimator.

Evaluation of target parameter of targeted MLE: Above we defined a template for the targeted MLE $P_n \rightarrow Q_n^*(P_n)$. One now evaluates the target $\Psi^F(Q_n^*(P_n))$ to obtain the wished estimator of $\psi_0 = \Psi^F(Q_0)$.

Statistical Inference:

Inference under the assumption that the treatment/censoring mechanism estimator converges to true g_0: Under this assumption, we can carry out influence curve based inference. In order to carry out statistical inference we can use the fact that

$$P_n D^*(Q_n^*(P_n), g_n(P_n)) = 0,$$

where $Q_n^*(P_n)$ is the targeted MLE. If D^* can be represented as an estimating function $D^*(\psi, Q, g)$ in ψ, then this corresponds with stating that $\psi_n^* = \Psi(Q_n^*)$ solves the estimating equation

$$P_n D^*(\psi_n^*, Q_n^*, g_n) = 0,$$

and statistical inference can now be based on the double robustness of the estimating function $P_0 D^*(\psi_0, Q, g) = 0$ if either $Q = Q_0$ or $g = g_0$. In particular, under the assumption that g_n converges to g_0, asymptotically conservative first order statistical inference can be based on the influence curve $D^*(\psi, Q, g_0)$ and corresponding confidence intervals $\psi_n^* \pm 1.96\sigma_n/\sqrt{n}$, where

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} D^*(\psi_n^*, Q_n^*, g_n)^2(O_i)$$

is an estimate of the variance of the influence curve. This follows from the fact that under regularity conditions, ψ_n^* is asymptotically linear with influence curve $D^*(\psi_0, Q, g_0)$ minus its projection on the tangent space of the model of g_n^*, where Q is the possibly miss-specified limit of Q_n^* (see van der Laan and Robins (2003)).

Collaborative double robust statistical inference: If we want to rely on double robustness, either due to high dimension of the treatment mechanism, or due to sparsity (w.r.t. target) so that the targeted maximum likelihood step is unstable, then we recommend the use of the collaborative targeted MLE mentioned above, which also involves the selection among targeted maximum likelihood estimators indexed by different candidate estimators of g_n and thereby different targeted MLE steps, where this choice is based on how much the clever covariates improve the fit of the log-likelihood (or loss-function specific risk) of the corresponding targeted maximum likelihood estimator. As shown in the Appendix of van der Laan and Gruber (2009), using this approach, one can still use the influence curve $D^*(Q, g)$ for statistical inference, with Q, g denoting the limits of Q_n^* and g_n, and one now relies on collaborative double robustness stating that g_n needs to converge to a true conditional distribution $g_0(Q)$, indexed by the limit Q of Q_n^*, that adjusts correctly for the residual bias due to misspecification of Q_n^* w.r.t Q_0. For details, we refer to van der Laan and Gruber (2009).

Off course, one can also use the bootstrap for statistical inference.

5 Application to marginal structural model for realistic individualized rules

For the sake of concreteness, let's consider the "When to start treatment" question in HIV research. Let time 0 denote the time at which the patient enrolls in the study. At this time, various measurements are made, including baseline CD4 count and viral load. Subsequently each patient is monitored at various times on a possibly fine discrete time scale (e.g., the time unit might represent a week), and followed up till end of follow up or death. Certain time-dependent variables might be recorded at regularly spaced monitoring times such as a viral load and CD4 count, or the virus might be sequenced. Other indicator variables of particular events such as life-status indicator, heart-attack indicator, cancer occurrence indicator, infection indicator, might be observed at irregularly spaced monotoring times. In that case, if such an event occurs, it is recorded, and one also knows at each time, if it has already occurred (i.e., if it has not been recorded, then it did not happen). At such intermediate events, certain time-dependent variables might be recorded, be-

yond the variables recorded at the planned monitoring times. We will refer to any such time as a monitoring time, but one will have to specify different types of monitoring times. Different types of monitoring times might result in the recording of different variables. Beyond these different type of monitoring times and the corresponding data collection at these time points, there is a time till death and time till right-censoring, both marking the end of follow up. The right-censoring event could be indexed by different types, such as right-censoring by the end of study or by a medical doctor's decision. The indicator process which jumps at the start of treatment is also observed. We are concerned with using n such independently and identically distributed longitudinal data structures to compare the efficacy of different treatment strategies. For the sake of illustration we will focus on targeted maximum likelihood estimation of the effect of different rules for when to start a patient on antiretroviral therapy.

The organization of this section is as follows. To start with we will discuss the format of the data on one patient, and the formulation of the likelihood of such a longitudinal data structure factorized according to the statistical graph defined by chronological time ordering. Subsequently, we will define various causal effects of "when to start rules" as parameters of interventions on this likelihood defined by the G-computation formula, including the unknown regression parameters in a marginal structural model for a family of realistic individualized "when to start treatment rules". We will then define a first stage super-learning maximum likelihood (or other loss-function based) estimator, and corresponding targeted maximum likelihood estimator of these unknown parameters.

5.1 The graph-factorized likelihood of the data structure for one unit:

The data on a subject will involve various lines of data, each line corresponding with a monitoring time (e.g., corresponding with an intermediate event time such as an infection/heart attack etc), as indicated by a time-stamp column, and a final line corresponding with the follow up time till time at analysis or till death or another event marking end of follow up. At each monitoring time, updates on a collection of time-dependent variables will be recorded, while the columns coding time-independent covariates remain (obviously) constant. The time-dependent variables that are not measured at

that monitoring time are either coded as missing or one imputes (e.g forward imputation) a value, and one creates an imputation indicator indicating if this measurement was imputed or an actual update. The final line of data provides the final time stamp, and either the censoring indicator column might jump to the value 1, and the type of censoring is coded, or, if this final time-stamp is time till death, then the life-status indicator jumps to the value 1.

Formally, we could code such a data set as follows. Let $W_j(t)$ be a time-dependent variable, defined as constant in between monitoring times, using forward imputation, and missing if no previous measurement is available, $j = 1, \ldots, J$. Let $\Delta_j(t)$ be an imputation indicator corresponding to $W_j(t)$, $j = 1, \ldots, J$. Let $N_j(t)$ be a counting process such as $R_j(t) = I(T_j \leq t)$ for time till event variables T_j, which are observed at all time points t, $j = 1, \ldots, J_1$. Updates of the variables $W(t)$ only occur at a time point for which at least one of the $dN_j(t) = 1$, i..e, at a time t at which one of the counting processes N_j jump. For example, if there is regular monitoring at fixed time points, beyond monitoring at random times, then one of the counting processes codes the regular monitoring times, while others code the random monitoring times. We can now define a process $O(t) = ((W_j(t), \Delta_j(t) : j), (N_j(t) : j))$, and we truncate this process at the minimum of end of follow up time and a maximal follow up time τ. We will suppress the missing indicators $\Delta_j(t)$ in notation below, and just refer to $W_j(t)$. The observed data structure for one unit is now given by $\bar{O} = (\bar{O}(t) : t \leq \tau)$. Note that in this file $\bar{O}$, a jump of N_j results in a new line of data.

If at a monitoring time stamp various measurements are made, then there might be additional time ordering within the time-stamp. For example, the time-stamp might correspond with a day, and it might be known that the medical doctor made a treatment decision that day based on a variety of newly recorded measures, and possibly additional measures were obtained that same day, after the treatment had been given. This kind of additional time ordering information is an important component of the causal graph necessary to obtain a valid G-computation formula respecting the time-ordering. In this case, for each t, $W(t)$ should be accompanied with a time-ordering vector making clear which groups of variables were measured at the same time and how these groups are ordered in time.

The factorization of likelihood respecting the time-ordering:

The likelihood of this longitudinal data structure $\bar{O}$ could be written as a product over time t starting at time 0. One starts with drawing the

baseline data at time 0. Subsequently, at each time point one draws from the conditional intensities of N_j, given past, including the planned monitoring intensities, time till event intensities, and end of follow up intensities such as time till death or time till right censoring intensities. As long as none of these events occur, one proceeds to the next time interval.

When one of these events occurs, then the possible additional random variables corresponding with that type of event are drawn, possibly sequentially according to an additional time ordering, thus always following the known time-ordering. One proceeds generating random variables like this, moving along in discrete time, till either death happens, another final event happens which marks the end of follow up, or the end of study is reached.

Such a likelihood $p_0(O) = \prod_t p_0(O(t) \mid \bar{O}(t-))$, can be factorized as

$$\begin{aligned} p_0(O) &= \prod_t \prod_j \lambda_j(t \mid Pa(dN_j(t)))^{dN_j(t)}(1 - \lambda_j(t \mid Pa(dN_j(t))))^{1-dN_j(t)} \\ & P(W(t) \mid Pa(W(t))), \end{aligned}$$

where we used an ordering for the components $dN_j(t)$ so that $Pa(dN_j(t)) = \bar{O}(t-), dN_1(t), \ldots, dN_{j-1}(t)$, and $Pa(W(t)) = \bar{O}(t-), dN(t)$. If $dN(t) = 0$ (i.e., each $dN_j(t) = 0$), then no events and monitoring occurs at time t so that $P(W(t) \mid Pa(W(t)))$ is degenerate.

Let $W(t) = (W^-(t), A(t), W^+(t))$, where $A(t)$ is a treatment decision at time t, $W^-(t)$ are variables recorded before $A(t)$, and $W^+(t)$ are variables recorded after t. This particular time-ordering at time t can depend on the realization of $dN(t)$: i.e. for different types of events, different variables might be collected, and for each such group of variables that recorded, we need to know what variables are pretreatment and post treatment decision at time t.

So we have, respecting the time-ordering,

$$P(W(t) \mid Pa(W(t))) = P(W^-(t) \mid Pa(W^-(t)))P(A(t) \mid Pa(A(t)))P(W^+(t) \mid Pa(W^+(t))).$$

The conditional probability distributions $P_{W^-(t)}$ and $P_{W^+(t)}$ can be factorized in terms of products of conditional densities of particular variables, and these conditional densities can be further factorized in terms of hazards of binary events. as we did in the previous sections.

We conclude that we have the following factorization of the likelihood in terms of conditional distribution of binary events, given the parent sets,

factored according to the known time-ordering and user supplied orderings in the case that there is no time ordering provided,

$$p_0(O) \;=\; \prod_t \prod_j P_{dN_j(t)} \prod_{jl} P_{W_{jl}^-(t)} \prod_{jl} P_{W_{jl}^+(t)} \prod_t g_{(A(t)}.$$

5.2 The likelihood factored in intervention mechanism and relevant factor.

Treatment mechanism: We can refer to $\prod_t g_{A(t)}$, with $g_{A(t)}(A(t) \mid Pa(A(t)))$, as the treatment mechanism, and it provides us also with a likelihood criterion that can be used to generate maximum likelihood estimators of this treatment mechanism. A special case of a process $A(t)$ is of the form $A(t) = I(S \leq t)$, where S is the time at which an antiretroviral therapy is started. In this case, $A(t)$ only jumps once.

If $dN(t) = 0$, then no treatment decisions are made at time t so that $g_{A(t)}$ is degenerate at such a parent set realization. If $dN(t) \neq 0$, then the treatment assignment mechanism can still depend heavily on the type of event that occurred: i.e., which $dN_j(t) = 1$. In particular, it might be the case that only for one type of event, treatment decisions are made, while for any of the other events coded by the counting processes $N_j(t)$, $A(t)$ will not be assigned/changed, so that it will still follow a degenerate probability distribution. Note that this treatment mechanism product over all times t reduces to a product of treatment probabilities at the finite (but random) time points for which $dN(t) \neq 0$ and for which there is experimentation in $A(t)$. For example, if N_1 jumps at the time point at which doctors generate measurements and make treatment decisions, and all other events as coded by N_j, $j = 2, \ldots, J$ do not generate treatment decisions with probability 1, then we have

$$g = \prod_t g_{A(t)} = \prod_{t:dN_1(t)=1} g_{A(t)}(\cdot \mid Pa(A(t)), dN_1(t) = 1).$$

Interventions on treatment: Interventions on this treatment assignment mechanism define interesting causal effects. However, it needs to be understood that such an intervention does not control the actual time points at which these treatment decisions can be enforced: i.e.. the time points at which $A(t)$ changes value are kept uncontrolled under an intervention on this

treatment mechanism, and, might be differential depending on the actual treatment intervention.

Interventions on treatment and timing of treatment changes: Therefore it is also of interest to intervene on both the treatment assignment mechanism as well as the monitoring mechanism that generates the monitoring times at which the treatment decisions can be enforced. For that purpose, suppose that N_1 is a counting process that jumps at times at which treatment decisions are made. It would now also be of interest to both intervene on the monitoring process N_1 as well as on the treatment decisions made at the monitoring times defined by N_1.

In that case the total mechanism defining the "treatment mechanism" is given by

$$g_1 \equiv \prod_t \lambda_{N_1(t)}(t)^{dN_1(t)}(1-\lambda_{N_1(t)}(t))^{1-dN_1(t)} g_{A(t)}.$$

The first factor concerns the assignment of monitoring times of type 1 and the second factor concerns treatment decisions at such times, but possibly also at monitoring times of different types. As a special case, one might have that treatment decisions are only made at monitoring times of type 1, so that intervening on N_1 and the treatment assignment mechanism at these monitoring times is an intervention on the complete treatment process. In this case, only $g_{A(t)}(A(t) \mid Pa(A(t)), dN_1(t) = 1)$ needs to be estimated, since conditioning on other realizations of $dN(t)$ makes $g_{A(t)}$ a degenerate distribution.

Right-censoring mechanism: We also define the factor that defines the likelihood for the right-censoring events. For example, if N_2 is the (only) counting process that codes right-censoring events that obstruct the complete observation of the outcome of interest Y, then we define the censoring mechanism as

$$g_2 = \prod_t \lambda_{dN_2(t)}(t)^{dN_2(t)}(1-\lambda_{dN_2(t)}(t))^{1-dN_2(t)}.$$

Combined set of conditional distributions we intervene upon: We will denote the combined treatment and censoring mechanism with $g = g_1 * g_2$. Here g_1 can be either the treatment mechanism or it can be both the treatment and monitoring mechanism, depending on the scientific question we wish to address.

Incorporation of causal graph knowledge beyond time-ordering: Causal graph knowledge can be incorporated by reducing the parent set of

the nodes, and or enforcing orderings of variables beyond the one implied by the time ordering. The time-ordering always has to be satisfied by any causal graph, so that any additional causal graph information provides additional ordering of all measured variables, and possible reductions of parent sets.

5.3 Target parameters of G-computation formula: Marginal structural model for intervention rules.

Causal effect of intervention on treatment decisions at uncontrolled monitoring times: Suppose that, in words, we wish to assess the mean outcome Y measured at a fixed time K since baseline under a dynamic treatment rule d_θ that start treatment right after a measured CD4 count falls below θ. Here Y might be defined as the indicator of still being alive at time K, or an absolute level of CD4 count, or a combination of death and CD4 count such as an indicator of death or CD4 is below a critical value.

We now need to decide how we can formally define this parameter as an operation on the time-ordered/causal graph factorized likelihood of the data. Firstly, let's consider the case in which we do not intervene on the monitoring process N_1 that generates the monitoring times at which biomarker data (e.g., CD4 and viral load) is generated and treatment decisions are made. In this case we only intervene on the treatment assignment rule at the monitoring times generated by the true intensity λ_1 of N_1. Such a rule might be that if the most recent measured CD4 count is below θ, then the anti-retroviral therapy is started.

Since our outcome Y is subject to right-censoring by some of the other intensities and we are only interested in the effect of the treatment intervention on the uncensored outcome, we also need to intervene on the censoring mechanism. Let A_2 be the censoring process which jumps from zero to one at a time point t in which a subject is right-censored by an event coded by one or more of the counting processes. Let $g = g_1 g_2$ denote the factor of the likelihood that generates the treatment A and right-censoring A_2 events.

To generate the counterfactual data under such a rule d_θ, one would now generate the data according to the likelihood as described above, but at a monitoring time t with $dN_1(t) = 1$ at which a treatment decision is possible, we would now apply the dynamic rule to set the indicator $A(t)$ of starting treatment at time t, and at each time at which censoring can occur, we set the relevant $dN_j(t) = 0$ for all j that code right-censoring events. Note that

this intervention would leave the monitoring process random as it is. So, if particular patients are badly monitored, even under a rule d_θ, they might start treatment at a much lower CD4 count than θ due to large time periods in which the patient's CD4 count is not observed.

Causal effect of intervention on monitoring times and treatment decisions.

If one is concerned that the monitoring process heavily affects the clinical outcome and one is concerned with extrapolation of the results to a population in which monitoring times are differently distributed, then one might wish to assess the effect of rules that intervene on both the monitoring mechanism, as well as the treatment assignment mechanism. For example, a rule might be that one monitors a patient every θ_1 months and thereby measures the CD4 count and viral load at these times, and that one starts anti-retroviral treatment when the CD4 count measured at that time is below θ_2. To generate the data under such a rule d_{θ_1,θ_2}, one would now generate the data according to the likelihood as described above, except one would not generate monitoring times based on λ_1 till one reaches $t = \theta_1$, at which time one sets the monitoring time at $t = \theta_1$, one draws from the conditional distribution of time-dependent covariates at that time point, conditional on being monitored at that time, and one assigns the when to start treatment decision according to the rule indexed by threshold θ_2. One proceeds over time following the data generation according to the likelihood, the set monitoring times, the set treatment starting rule, and right-censoring set at infinity. It is also of interest to consider an intervention on monitoring that corresponds with randomly drawing monitoring times from a user supplied monitoring mechanism, but we will not consider this case below.

Identifiability of target parameter: These rules have to be realistic rules in order to make the corresponding counterfactual probability distributions identifiable from the observed data. For example, the intensity λ_1 needs to have support on these regularly spaced monitoring times $k\theta_1$, $k = 1, 2, \ldots$, and the lower the support the harder it will be to reliably estimate the counterfactual distribution for that choice of θ_1. If in the actual study people were regularly followed, say every 3 months, and that variations on the monitoring times were at most one month off, then one would expect a good support for $\theta \in [3 - \delta, 3 + \delta]$ for an appropriately chosen δ. Similarly, the when to start treatment decision rule needs to be supported by the medical doctors that made these decisions in the actual study. For example, if the medical

community supports the starting of the antiretroviral therapy at CD4 counts between 200 and 400, in the sense that there is experimentation across that range, then one should choose θ_2 in that range. Finally, the right-censoring intensity should never equal 1, whatever the history, up till the time point K at which the outcome can be measured. Thus, if K is selected too large, then the latter assumption might become practically violated. In addition, if there are events that imply right-censoring at the next time point with probability 1, then one might need to include such events in the definition of the outcome Y, so that the effect of treatment on that Y is still identifiable from the data, and such effects will then need to be honestly interpreted.

These counterfactuals Y_θ indexed by the rule d_θ for the treatment process $\bar{A}$, either only including the treatment decisions or also including the monitoring time process N_1, are now defined by its probability distribution Q_θ define above as an intervention on the the graph-factored likelihood, obtained by excluding the factors $g = g_1 g_2$ that are set by the rule, and setting the values of treatment $\bar{A}$ and right-censoring $\bar{A}_2$ according to the rule d_θ in any of the conditioning events of the other factors of the likelihood.

We can now define the parameter of interest as a projection of EY_θ onto a working model, thereby creating smoothing parameters of the complete response curve $\theta \to EY_\theta$ for which larger data support is available so that it can be estimated using semi-parametric model efficiency theory and methodology. That is, EY_θ is often not path-wise differentiable, while such a summary parameter will be path-wise differentiable.

Marginal structural models for realistic individualized treatment rules: Specifically, let $m_\beta(\theta, V)$ be a working model for the conditional mean of Y_θ, under rule d_θ (controlling treatment and censoring), given a baseline (for example) CD4 count V, such as

$$m_\beta(\theta, V) = \beta_0 + \beta_1\theta + \beta_2\theta^2 + \beta_3\theta V + \beta_4\theta^2 V.$$

We now define the target parameter as

$$\Psi(P_0) = \Psi^F(Q_0) = \arg\min_\beta E_{Q_0} \sum_\theta h(\theta, V)(Y_\theta - m_\beta(\theta, V))^2,$$

where we remind the reader that Q_0 denotes the factor of the probability distribution of O: $p_0 = Q_0 g$. Here $h(\theta, V)$ is a user supplied weight function. Thus Ψ is only a parameter of P_0 through its Q_0-factor. We refer to m_β as a marginal structural (working) model for realistic treatment rules (van der Laan and Petersen (2007)).

Evaluation of target parameter: Given such an estimator Q_n or its targeted MLE update Q_n^* defined below, the evaluation of $\Psi(Q_n)$ can be based on Monte-Carlo simulation. One first samples a larger number B of observations $Y_{\theta,b}, V_b$, $b = 1, \ldots, B$ for each possible θ, from the corresponding G-computation formula of the distribution of Y_θ. Then,

$$\Psi^F(Q_n) = \arg\min_\beta \sum_\theta \frac{1}{N} \sum_{b=1}^{B} h(\theta, V_b)(Y_{\theta,b} - m_\beta(\theta, V_b))^2.$$

In other words, we simply replace the expectation E_{Q_0} of a function $f_{\theta,\beta}(Y_\theta, V)$ of Y_θ, V in the definition of $\Psi^F(Q_0)$ by an expectation w.r.t. empirical distribution of the B draws $(Y_{\theta,b}, V_b)$, $b = 1, \ldots, B$.

5.4 Adaptive Maximum Likelihood Estimation: Super learning.

Given an estimator Q_n of Q_0, one obtains the estimator $\Psi(Q_n)$ of $\psi_0 = \Psi^F(Q_0)$. Since the likelihood factor Q factorizes in a product over conditional distributions of binary factors, we can estimate this with the loss-based super-learning methodology. Thus one applies super learners for binary predictions, possibly pooled across many of the binary predictions, pooling across time and or across different levels of ordered variables indexing the binary variables. The overall log-likelihood or pooled weighted squared error loss function for Q_0 could be employed for fine tuning the choice and degree of pooling, only considering sensible pooling strategies. Super-learning could then be based on a library of algorithms for estimation of the complete Q_0 based on this overall loss function for Q_0. Some of the candidate algorithms might involve super-learning itself of binary predictors possibly using different pooling strategies.

To be specific, let's consider modeling the conditional distribution of CD4 count at time t, given its parents. We will model this in terms of conditional binary distributions. Suppose that we can view CD4 as an ordered discrete variable with levels $l = 1, \ldots, L$, possibly defined by the L equally spaced quantiles of the marginal empirical distribution of CD4 counts. Let $Q_{CD4(t)}$ denote the conditional distribution of $CD4$ count at time t, conditional on the parent nodes and that the person is monitored at time t so that the CD4 count process is at risk of changing. We write this conditional probability

distribution of $CD4(t)$ in terms of binary conditional distributions

$$Q_{CD4(t)}(CD4(t)) = \prod_{j=1}^{L} Q_{I(CD4(t)=j)}(1)^{CD4(t)=j} Q_{I(CD4(t)=j)}(0)^{CD4(t)\neq j}.$$

If we define the discrete hazard

$$\lambda_{CD4(t)}(j) \equiv P(CD4(t) = j \mid CD4(t) \geq j, Pa(CD4(t))$$

of CD4 count at time t, then it follows that

$$Q_{CD4(t)}(l) = \prod_{j=1}^{l-1}(1 - \lambda_{CD4(t)}(j))\lambda_{CD4(t)}(l).$$

The likelihood for this discrete hazard is thus given by

$$L(\lambda_{CD4(t)}) = \prod_{i=1,Mon_i(t)=1}^{n} \prod_{l=1}^{CD4_i(t)-1} (1 - \lambda_{CD4(t),i}(l))\lambda_{CD4(t),i}(CD4_i(t)).$$

The likelihood of this discrete hazard $\lambda_{CD4} \equiv (\lambda_{CD4(t)}(j) : t, j)$ viewed as a function in both time t and the CD4 count level j is thus given by

$$L(\lambda_{CD4}) = \prod_{t} \prod_{i=1,Mon_i(t)=1}^{n} \prod_{i=1,Mon_i(t)=1}^{n} \prod_{l=1}^{CD4_i(t)-1} (1-\lambda_{CD4(t),i}(l))\lambda_{CD4(t),i}(CD4_i(t)).$$

One can now carry out estimation of this nonparametric function λ_{CD4} based on this log-likelihood loss function. In particular, one can apply super-learning based on this loss function.

We note that if monitoring is really random, then one will have few subjects that have a monitoring time at a given time point t within a fine grid of time points. As a consequence, in that case the t-specific likelihood for $\lambda_{CD4(t)}$ provides too little information for estimation of $\lambda_{CD4(t)}$. Thus, it will be essential to use the combined likelihood pooling across time the time points t provided above.

5.5 Calculation of least favorable model for targeted MLE step.

We now focus our attention on the definition of the fluctuation function required to carry out the targeted MLE step.

Inverse Probability of Censoring Weighted Function: Let X represent the collection of action specific counterfactuals controlling the intervention nodes defined by $\bar{A}_1, \bar{A}_2$. We will first define an IPCW-estimating function of O for the parameter EY_θ for a given θ, before presenting the IPCW-function of O for the MSM-parameter Ψ.

We have

$$EY_\theta = E\left(\frac{I(\bar{A}_1 = d_\theta(L), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} Y\right),$$

where $g = g_1 g_2$ represents the product of conditional distributions of the intervention nodes $\bar{A}_1$ and $\bar{A}_2$. Thus an IPCW-estimating function of EY_θ is given by (see also van der Laan and Petersen (2007))

$$\frac{I(\bar{A}_1 = d_\theta(L)), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} (Y - EY_\theta).$$

By a similar argument, it follows that the IPCW-estimating function of $\Psi(Q_0)$ is given by

$$D_{IPCW}(O) = \sum_\theta h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) \frac{I(\bar{A}_1 = d_\theta(L), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} (Y - m_{\psi_0}(\theta, V)).$$

For example, if we also intervene on the monitoring times at which treatment can be changed, then $g(\bar{A} \mid X)$ involves a product over time of the likelihood of monitoring events and a treatment event if monitoring occurred, and no censoring event, always conditioning on the parent sets implied by the graph implied by time ordering and possibly additional causal graph assumptions.

The efficient influence curve and corresponding clever covariates for binary factors: We first note that the IPCW-estimating function can be represented as

$$D_{IPCW}(O) = \frac{1}{g(\bar{A}|X)} \sum_\theta h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) I(\bar{A}_1 = d_\theta(O), \bar{A}_2(K) = 0)(Y - m_{\psi_0}(\theta, V))$$
$$\equiv \frac{D_1(O)}{g(\bar{A}|X)}.$$

We have that the Q_0-factor of the density of the data is represented as

$$Q_0 \;=\; Q_{W(0)} \prod_t \prod_j Q_{dN_j(t)} \prod_{jl} Q_{W_{jl}^-(t)} \prod_{jl} Q_{W_{jl}^+(t)}.$$

Here we exclude the conditional distributions of the counting processes corresponding with right-censoring and or monitoring events, depending on how

$\bar{A}$ is defined. This Q_0-factor identifies the G-computation formula for the distribution of the data under the individualized interventions d_θ.

Let T_Q be the tangent space of Q_0 at $P_{Q,g}$. This tangent space T_Q can be decomposed orthogonally as:

$$T_Q = T_{Q_{W(0)}} + \sum_{tj} T_{\lambda_j,t} + \sum_{tjl} T_{Q^-_{tjl}} + \sum_{tjl} T_{Q^+_{tjl}},$$

where $T_{Q_{W(0)}}$ is the tangent space of the marginal probability distribution of $W(0)$, $T_{\lambda_j,t}$ is the tangent space of the j-th intensity $\lambda_j(t)$, $T_{Q^-_{tjl}}$ is the tangent space of conditional distribution of $W^-_{jl}(t)$, and $T_{Q^+_{tjl}}$ is the tangent space of conditional distribution of $W^+_{jl}(t)$.

By our general Theorem, we have that the efficient influence curve $D^* = \Pi(D \mid T_Q)$ can be represented as

$$D^* = \Pi(D \mid T_Q) = D_0 + \sum_{tj} D_{t,\lambda_j} + \sum_{tjl} D^-_{tjl} + \sum_{tjl} D^+_{tjl},$$

where $D_0 = E(D^* \mid W(0))$, and for each other factor, we can represent it as $D_{\lambda_j,t} = C_{tj}(dN_j(t) - \lambda_j(t))$, $D^-_{tjl} = C^-_{tjl}(W^-_{jl}(t) - Q^-_{tjl}(1))$, $D^+_{tjl} = C^+_{tjl}(W^+_{jl}(t) - Q^+_{tjl}(1))$, where these (clever covariates) $C_{tj}, C^+_{tjl}, C^-_{tjl}$ are defined below. These functions D_{t,λ_j}, D^+_{tjl} and D^-_{tjl} are zero at corresponding parent histories which deterministically predict the value of the corresponding binary variable $dN_j(t)$, $W^+_{jl}(t)$, $W^-_{jl}(t)$, respectively. For example, if $W^-_{jl}(t)$ is only generated at a monitoring time t generated by N_1, then $\sum_t D^-_{tjl} = \sum_{t:dN_1(t)=1} D^-_{tjl}$ reduces to a sum at the random monitoring times at which $dN_1(t) = 1$. On the other hand, if N_1 is at risk of jumping at any of the time points t, then $\sum_t D_{t,\lambda_1}$ remains a sum over all time points t.

For each binary variable $dN_j(t)$, $W^+_{jl}(t)$, $W^-_{jl}(t)$ we define $\bar{A}(t,j)$, $\bar{A}^+(t,j,l)$, $\bar{A}^-(t,j,l)$, respectively, as the A-nodes that are included in the parent set of that binary variable. For each binary variable $dN_j(t)$, $W^+_{jl}(t)$, $W^-_{jl}(t)$ we define $\overline{a}(t,j)$, $\overline{a}^+(t,j,l)$, $\overline{a}^-(t,j,l)$ as the future path, corresponding with a complete path $\bar{a}$, starting right after where $\bar{A}(t,j)$, $\bar{A}^+(t,j,l)$, $\bar{A}^-(t,j,l)$ stops.

Suppose that the process A we control includes the monitoring process N_1. In that case, $\bar{A}(t,j) = \bar{A}(t-), dN_1(t), dA_2(t))$ includes all actions up till and including previous time point and the censoring and monitoring event at time t, $\bar{A}^-(t,j,l)$ equals $\bar{A}(t-), dN_1(t), dA_2(t)$ as well, and $\bar{A}^+(t,j,l)$ equals

$\bar{A}(t-), dN_1(t), dA_1(t), dA_2(t)$ including now also the treatment decision at time t. If our process A does not control monitoring N_1, then exclude N_1 from the statement in the previous sentence.

The formulas for C_{tj}, C^+_{tjl}, and C^-_{tjl} can now be defined as

$$
\begin{aligned}
&C_{tj} = \tfrac{1}{g(\bar{A}(t,j)|X)} \times \\
&\{E_Q(\textstyle\sum_{\overline{a}(t,j)} D_1 \mid dN_j(t) = 1, Pa(dN_j(t))) - E_Q(\sum_{\overline{a}(t,j)} D_1 \mid dN_j(t) = 0, Pa(dN_j(t)))\} \\
&C^+_{tjl} = \tfrac{1}{g(\bar{A}(t,j,l)^+|X)} \times \\
&\{E_Q(\textstyle\sum_{\overline{a}(t,j,l)^+} D_1 \mid W^+_{jl}(t) = 1, Pa(W^+_{jl}(t))) - E_Q(\sum_{\overline{a}(t,j,l)^+} D_1 \mid W^+_{jl}(t) = 0, Pa(W^+_{jl}(t))\} \\
&C^-_{tjl} = \tfrac{1}{g(\bar{A}(t,j,l)^-|X)} \times \\
&\{E_Q(\textstyle\sum_{\overline{a}(t,j,l)^-} D_1 \mid W^-_{jl}(t) = 1, Pa(W^-_{jl}(t))) - E_Q(\sum_{\overline{a}(t,j,l)^-} D_1 \mid W^-_{jl}(t) = 0, Pa(W^-_{jl}(t)))\}
\end{aligned}
$$

Here we used short-hand notation for $\sum_{\overline{a}(t,j)} D_1(O_{\bar{A}(t,j),\overline{a}(t,j)})$, and similarly for the other two terms. If the parent set implies that there is no experimentation in the node, then this clever covariate is not defined, and is also never needed, as stated above. If one wants to extend the definition, then one could simply define the clever covariate as zero for any parent set in which there is no experimentation in the binary node.

For each binary node, the clever covariate can also be represented as the difference of the conditional expectation of $D_{IPCW} = D_1/g(\bar{A} \mid X)$ given the binary node equals 1, and the conditional expectation of D_{IPCW} given the binary node equals zero, and in both cases one also conditions on the parent set of the binary node. In other words, it is a choice to either integrate out over the future sample paths $\overline{a}$ so that the clever covariate factors in a g and Q-factor, or not. One can evaluate these prediction representations of the clever covariate by Monte-Carlo simulation which involves drawing from all future factors of the density (either from Q only, or, from both Q and g, depending on the representation), starting at which the parent set left-off.

As in the main paper, these clever covariates provides us with the least favorable model through Q_n (at $\epsilon = 0$) with fluctuation parameter ϵ whose score at $\epsilon = 0$ equals the efficient influence curve $D^*(Q_n, g_n)$, and defines the corresponding targeted MLE Q^*_n and $\Psi(Q^*_n)$.

5.6 Targeted maximum likelihood estimation at degenerate initial estimator of intermediate conditional distribution.

For simplicity, let's consider the case that our initial estimator provides deterministic predictions for any of the intermediate time-dependent covariates, so that the clever covariates for all intermediate factors equals zero. As a consequence, the targeted MLE only involves updating the conditional distribution of the final node Y, given its parents.

Consider the IPCW-estimating function:

$$D_{IPCW}(O) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) \frac{I(\bar{A}_1 = d_\theta(O), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} (Y - m_{\psi_0}(\theta, V)).$$

Under a degenerate distribution for all intermediate variables, we only have to project this onto the tangent space of the distribution of $L(0)$ and the conditional distribution of Y, given $\bar{L}(K), \bar{A}(K)$. We will now present this efficient influence curve at such a Q, and, for the sake of illustration, we will show that it represents an unbiased estimating function of ψ at a correctly specified g_0, and arbitrarily misspecified Q. The efficient influence curve has now only two components we will denote with $D_1^*(Q, g_0)$ and $D_2^*(Q, g_0)$ respectively. We have

$$D_1(Q, \psi_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) (E_Q(Y_\theta \mid L(0)) - m_{\psi_0}(\theta, V))$$

and, using $d_{\theta,0}$ to denote both the treatment rule and the no-censoring intervention,

$$D_2(Q, g_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) \frac{I(\bar{A} = d_{\theta,0}(O))}{g(\bar{A} \mid X)} (Y - E_Q(Y \mid \bar{A}(K), \bar{L}(K))).$$

Thus the clever covariate we add to an initial estimator of the conditional distribution of Y, given $\bar{A}(K), \bar{L}(K)$, is given by

$$C(g_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) \frac{I(\bar{A} = d_{\theta,0}(O))}{g_0(\bar{A} \mid X)}.$$

Let Q_n^* be the targeted MLE based on an initial estimator Q_n defined by the empirical distribution $Q_{n,L(0)}$ of $L(0)$ and an initial estimator $Q_{n,Y}$ of Y, given

$\bar{A}(K), \bar{L}(K)$, and degenerate distributions $Q_{n,d}$ for all the conditional distributions of intermediate variables/time-dependent confounders. The targeted maximum likelihood estimator solves the efficient influence curve equation $P_n D^*(Q_n^*, g_n, \Psi(Q_n^*)) = 0$.

We now show that indeed, as predicted by the double robustness of the efficient influence curve, $P_0 D^*(Q, g_0, \Psi(Q)) = 0$ implies $\Psi(Q) = \Psi(Q_0)$, showing that the targeted MLE $\Psi(Q_n^*)$ solves an unbiased estimating function in ψ at a correctly specified g_0. Note, by first conditioning on X in $P_0 D_2$,

$$\begin{aligned} P_0 D^*(Q, g_0, \Psi(Q)) &= P_0 D_1(Q, \Psi(Q)) + P_0 D_2(Q, g_0) \\ &= E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m_\psi(\theta, V)(E_Q(Y_\theta \mid L(0)) - m_\psi(\theta, V)) \\ + E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m_\psi(\theta, V)(Y_\theta - E_Q(Y \mid \bar{A}(K) = d_\theta(\bar{L}(K), \bar{L}(K))). \end{aligned}$$

Now, note that, under the degenerate distribution Q, we have

$$E_Q(Y \mid \bar{A}(K) = d_\theta(\bar{L}(K), \bar{L}(K))) = E_Q(Y_\theta \mid L(0)).$$

Thus, we have

$$P_0 D^*(Q, g_0, \Psi(Q)) = E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m_\psi(\theta, V)(Y_\theta - m_{\Psi(Q)}(\theta, V)).$$

By definition of $\Psi(Q_0)$ we have that the latter equation equals zero at ψ_0. Thus, we can conclude that indeed, under a weak identifiability condition on the working model m_ψ, $P_0 D^*(Q, g_0, \Psi(Q)) = 0$ implies $\Psi(Q) = \psi_0 = \Psi(Q_0)$.

This targeted MLE only involves adding a clever covariate to $Q_{n,Y}$ and doing a single step update. This updated distribution, only updating $Q_{n,Y}$, equals the targeted MLE at such an initial Q_n.

Discussion on using degenerate fits to simplify targeted MLE: Even though the degeneracy of the initial estimator results in a simple to compute T-MLE, it is questionable till what degree this should be an issue to consider. Given available software, the actual practical performance will be the driving force in such a decision over time. We can simplify the T-MLE in less dramatic ways, by enforcing the degeneracy for most time-dependent variables, but truly modelling the conditional probability distribution for the most important time-dependent confounders. This might result in a highly

efficient T-MLE, while it still only involves updating the non-degenerate conditional distributions. In particular, such a T-MLE can still be obtained in closed form by using our backwards solving algorithm and using a separate updating step (i.e. variation independent fluctuation parameters for the conditional distributions) for these non-degenerate conditional distributions.

Specifically, in the when to start treatment application, it is well known that CD4 count and viral load are the most important time-dependent confounders. In addition, we consider dynamic rules for when to start the treatment responding to these time-dependent confounders. Thus, in this case, it seems particularly appropriate to estimate the actual conditional distributions of CD4 and viral load at the intermediate monitoring times.

The clever covariates of the CD4 count or viral load involve the evaluation of conditional expectations of the future outcome given the parent nodes. Since these conditional expectations are calculated under a Q-fit that uses a deterministic system for all variables except viral load, CD4, and the final outcome Y, this only involves random generation of the future CD4 counts, viral loads, and final Y. All other nodes are generated deterministically according to a fitted prediction function.

We also note that the degenerate conditional distributions do not need to be factored in terms of binary conditional distributions since these will not be updated anyway. Instead, one could simply predict the mean outcome for any continuous or ordered categorical variable from its parent nodes, and put probability 1 on that predicted value.

5.7 Dimension reduction.

We refer to our subsection on dimension reduction which shows that one can reduce the dimension of the time-dependent process $L(t)$ to few univariate time dependent processes, beyond the time-dependent covariates used in the rule d_θ, at little loss of information, while still using the same fit of the g-factor adjusting for all relevant variables.

6 Application to causal effect of point treatment, allowing for right-censoring and utilizing time-dependent covariates.

In this section, we consider a simplified version of the general data structure covered in previous section. Suppose we observe

$$O = (W, A, A_2(1), L(1), \ldots, A_2(K), L(K), A_2(K+1), Y),$$

where W is baseline covariates, A is a treatment assigned at baseline, $A_2(j) = I(C \leq j)$ is the indicator of being right-censored at time j, C is right-censoring time, $L(j)$ is the biomarker (e.g., CD4 count) measured at time j, using forward imputation if $C \leq j$, and Y is the final outcome of interest, but affected by right-censoring. For example, $Y = I(T \leq K+1)(1 - A_2(K+1))$ is the indicator of a time T till failure at time $K+1$ and not being censored. We assume $L(j)$ is ordered and discrete values with values $m = 1, \ldots, M$. The treatment A could be randomized as in a randomized controlled trial.

Let $\psi_0 = EY(1) - Y(0)$ be the additive causal effect of the binary treatment, where $Y(a)$ is defined as the random variable with probability distribution defined by the G-computation formula under the intervention $A = a$ and $C = \infty$ (i.e., no censoring), $a \in \{0, 1\}$.

The G-computation formula requires estimation of the marginal distribution of W, the conditional distribution of $L(j)$, given past and not being right-censored, $j = 1, \ldots, K$, and the conditional distribution of Y, given past and not being right-censored. The marginal distribution is estimated with the empirical distribution of W. If Y is binary, the conditional distribution of Y, given past and not being right-censored, is estimated with loss-based super-learning based on the log-likelihood loss function, and, if Y is continuous, we estimate the conditional mean of Y with loss-based super-learning based on the squared error loss function: the efficient influence curve of ψ_0 only depends on the conditional distribution of Y through its conditional mean. The conditional distribution of $L(j)$, given past, and not being right-censored, is also estimated with loss-based (super) learning using the log-likelihood loss function, but we will first factorize the conditional density as

$$P(L(j) = l \mid \cdot) = \prod_{m \leq l-1} (1 - P(L(j) = m \mid L(j) \geq m, \cdot)) P(L(j) = l \mid L(j) \geq l, \cdot).$$

In other words, we code $L(j)$ as a a vector of binaries $I(L(j)=1), I(L(j)=2), \ldots, I(L(j)=M)$, and, factor the likelihood of $L(j)$ accordingly. Thus, it remains to estimate the conditional hazard $P(L(j)=m \mid L(j) \geq m, Pa(L(j)))$. We could estimate this with super-learning smoothing in both time j and level m.

The G-computation formula for the counterfactual distribution of $W, \bar{L}, Y$ under intervention $A=a$ and no-censoring $\bar{A}_2 = 0$ is given by

$$\begin{array}{l} P_a(W, L(1), \ldots, L(K), Y) = \\ Q_W(W) \prod_{j=1}^{K} \prod_m Q_{L(j,m)}(L(j,m) \mid Pa(L(j,m), A=a, \bar{A}_2(j)=0) \\ Q_Y(Y \mid Pa(Y), A=a, \bar{A}_2(K)=0). \end{array}$$

The conditional distribution of $L(j,m) = I(L(j)=m)$, given its parents $(L(j,1), \ldots, L(j,m-1)), Pa(L(j))$ is degenerate if one of the indicators $L(j,l)$ with $l \leq m-1$ is already equal to 1.

The targeted maximum likelihood step now involves adding clever covariates to the logistic regression fits of the conditional distributions of $L(j,m)$, $j=1,\ldots,K$, $m=1,\ldots,M$, and the logistic or normal error (i.e., least squares, if Y is continuous) regression fit of Y. These clever covariates for the conditional distribution of the binary $L(j,m)$ are

$$\begin{array}{rcl} C_{jm}(Q,g) & = & E_{Q,g}(D_{IPCW} \mid L(j,m)=1, Pa(L(j,m)), \bar{A}_2(j)=0) \\ & & -E_{Q,g}(D_{IPCW} \mid L(j,m)=0, Pa(L(j,m)), \bar{A}_2(j)=0), \end{array}$$

where

$$D_{IPCW}(O) = Y \left\{ \frac{I(A=1, \bar{A}_2=0)}{g(A, \bar{A}_2 \mid X)} - \frac{I(A=0, \bar{A}_2=0)}{g(A, \bar{A}_2 \mid X)} \right\}.$$

So, calculation of the clever covariates requires, for each subject i, for each time j with $C_i > j$, and each m with $L_i(j) \geq m$, Monte-Carlo simulation to evaluate the conditional mean of D_{IPCW}, conditional on $Pa_i(L_i(j,m))$ for which $L_i(j) \geq m$ and $C_i > j$. This corresponds with imputing a $Y(1)$ and $Y(0)$, and thereby a $Y(1)-Y(0)$, for each subject, based on history of that subject at time j, across j.

The clever covariate to fluctuate the conditional distribution of Y is given by

$$C_Y = \left\{ \frac{I(A=1, \bar{A}_2=0)}{g(A, \bar{A}_2 \mid X)} - \frac{I(A=0, \bar{A}_2=0)}{g(A, \bar{A}_2 \mid X)} \right\}.$$

If Y is binary, the ϵC_Y is added on the logistic scale, and if Y is continuous, one adds ϵC_Y to the fitted conditional mean of Y.

The iterative targeted maximum likelihood algorithm can now be applied to obtain the targeted maximum likelihood estimator Q_n^* and corresponding $\Psi(Q_n^*)$. Statistical inference can be based on the estimated influence curve $D^*(Q_n^*, g_n)$. If one assumes that g_n converges to g_0, faster than Q_n^* converges to Q_0, then it makes sense to use as loss function for Q $L(Q) = D^*(Q, g_n)^2$, and thereby using loss-function based cross-validation to select among different targeted maximum likelihood estimators Q_n^*. In this manner, one is guaranteed to asymptotically select the targeted maximum likelihood estimator that results in the most efficient estimator of ψ_0.

7 Discussion

As mentioned in our template, in van der Laan and Gruber (2009) we present a variety of proposals for generating a sequence of targeted MLE's Q_j^* coupled with a treatment mechanism estimator g_j, each defined as the result of the targeted MLE algorithm that maps an initial estimator Q_j and treatment mechanism estimator g_j into an targeted MLE update Q_j^*, indexed by treatment/censoring mechanism estimators g_j that are increasingly nonparametric in j, and "initial" Q_j that themselves might represent a targeted MLE update of a previous initial estimator.

Such a sequence of candidate targeted MLE's is constructed to be increasing in the empirical risk of a loss function for Q_0 (e.g., log-likelihood), and corresponds with increasing levels of targeted bias reduction (since the later ones use an estimator of the treatment mechanism g_0 that is more nonparametrically estimating g_0 than estimator used in previous, and using g_0 results in the full bias reduction for the target parameter). Given such a constructed sequence of candidate targeted MLE's, one now selects the index of this sequence with the minimizer of the cross-validated risk of the loss function (see van der Laan and Gruber (2009)).

The main idea of collaborative targeted MLE is that targeted maximum likelihood estimators of ψ_0 are defined by an estimator Q_n^* of Q_0, so that a loss function (i.e. empirical criterion) can be used to evaluate different targeted maximum likelihood estimators that only differ in different degrees of targeted bias reduction. In this manner we can fine tune the bias reduction. For example, this adaptive selection guarantees that the targeted maximum

likelihood step (i.e., the choice of g_j) is actually improving the fit of Q_n^* w.r.t. the loss function, thereby dealing with the possible problem that the choice of g_j actually deteriorates the estimator relative to the initial estimator.

This selection approach for selection among candidate estimators g_j is not only theoretically grounded by oracle properties of the cross-validation selector, but also by the collaborative double robustness of the efficient influence curve as proved in van der Laan and Gruber (2009). This collaborative double robustness shows that the bias reduction for ψ_0 is achieved by an estimator g_n that correctly adjusts for the covariates that are still helpful in improving the fit of Q_n (i.e., the covariates that deal with the residual bias taking into account the initial estimator), but that covariates that are not needed to remove bias w.r.t. ψ_0 can be ignored. The fine-tuning of the bias reduction of the targeted maximum likelihood estimator through the collaborative targeted maximum likelihood estimator can, and typically should, be applied to further improve the finite sample mean squared error of the resulting estimator of ψ_0.

Onr purpose of the current paper is to lay the ground work for the software implementation of targeted maximum likelihood estimators that also incorporate time-dependent covariates. Time dependent covariates allow reasonably accurate imputations of the clinical outcome based on recent history, thereby allowing for significant potential gains in both efficiency and bias (see formula of effcicient influence curve/clever covariates). Even though almost all current clinical trials collect time dependent covariates, these important sources of information have been ignored. This provides one important application of the targeted maximum likelihood estimator, as presented in this paper, to be investigated. Other important applications are the assessment of causal effects of treatment rules in sequentially randomized controlled trials, and observational studies.

References

A. Abadie and G.W. Imbens. Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74:235–67, 2006.

P.K. Andersen, O. Borgan, R.D. Gill, and N. Keiding. *Statistical Models Based on Counting Processes*. Springer-Verlag New York, 1993.

O. Bembom, M.L. Petersen, S.-Y. Rhee, W. J. Fessel, S.E. Sinisi, R.W.

Shafer, and M.J. van der Laan. Biomarker discovery using targeted maximum likelihood estimation: Application to the treatment of antiretroviral resistant hiv infection. *Statistics in Medicine*, page http://www3.interscience.wiley.com/journal/121422393/abstract, 2008.

O. Bembom, M.J. van der Laan, T. Haight, and I.B. Tager. Lifetime and current leisure time physical activity and all-cause mortality in an elderly cohort. *Epidemiology*, 2009.

Oliver Bembom and Mark van der Laan. Statistical methods for analyzing sequentially randomized trials, commentary on jnci article adaptive therapy for androgen independent prostate cancer: A randomized selection trial including four regimens, by peter f. thall, c. logothetis, c. pagliaro, s. wen, m.a. brown, d. williams, r. millikan (2007). *Journal of the National Cancer Institute*, 99(21):1577–1582, 2007.

P.J. Bickel, C.A.J. Klaassen, Y. Ritov, and J. Wellner. *Efficient and Adaptive Estimation for Semiparametric Models*. Springer-Verlag, 1997.

J. Bryan, Z. Yu, and M.J. van der Laan. Analysis of longitudinal marginal structural models. *Biostatistics*, 5(3):361–380, 2003.

S. Dudoit and M.J. van der Laan. Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. *Statistical Methodology*, 2(2):131–154, 2005.

R. Gill and J.M. Robins. Causal inference in complex longitudinal studies: continuous case. *Ann. Stat.*, 29(6), 2001.

R.D. Gill, M.J. van der Laan, and J.M. Robins. Coarsening at random: characterizations, conjectures and counter-examples. In D.Y. Lin and T.R. Fleming, editors, *Proceedings of the First Seattle Symposium in Biostatistics*, pages 255–94, New York, 1997. Springer Verlag.

D.F. Heitjan and D.B. Rubin. Ignorability and coarse data. *Annals of statistics*, 19(4):2244–2253, December 1991.

M. A. Hernan, B. Brumback, and J. M. Robins. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11(5):561–570, 2000.

M. Jacobsen and N. Keiding. Coarsening at random in general sample spaces and random censoring in continuous time. *Annals of Statistics*, 23:774–86, 1995.

S. Keleş, M. van der Laan, and S. Dudoit. Asymptotically optimal model selection method for regression on censored outcomes. *Technical Report, Division of Biostatistics, UC Berkeley*, 2002.

P. Lavori and R. Dawson. A design for testing clinical strategies: biased adaptive within-subject randomization. *Journal of the Royal Statistical Society, Series A*, 163:2938, 2000.

P. Lavori and R. Dawson. Dynamic treatment regimes: practical design considerations. *Clinical trials*, 1:920, 2004.

K.L. Moore and M.J. van der Laan. Covariate adjustment in randomized trials with binary outcomes. Technical report 215, Division of Biostatistics, University of California, Berkeley, April 2007.

K.L. Moore and M.J. van der Laan. Application of time-to-event methods in the assessment of safety in clinical trials. In Karl E. Peace, editor, *in Design, Summarization, Analysis & Interpretation of Clinical Trials with Time-to-Event Endpoints*. Chapman and Hall, 2009.

S. Murphy. An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*, 24:14551481, 2005.

S.A. Murphy. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B*, 65(2):?, 2003.

S.A. Murphy, M.J. van der Laan, and J.M. Robins. Marginal mean models for dynamic treatment regimens. *Journal of the American Statistical Association*, 96:1410–1424, 2001.

R. Neugebauer and M.J. van der Laan. Why prefer double robust estimates. *Journal of Statistical Planning and Inference*, 129(1-2):405–426, 2005.

J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2000.

Maya L. Petersen, Steven G. Deeks, Jeffrey N. Martin, and Mark J. van der Laan. History-adjusted marginal structural models: Time-varying effect modification and dynamic treatment regimens. Technical report 199, Division of Biostatistics, University of California, Berkeley, December 2005.

E.C. Polley and M.J. van der Laan. Predicting optimal treatment assignment based on prognostic factors in cancer patients. In Karl E. Peace, editor, *in Design, Summarization, Analysis & Interpretation of Clinical Trials with Time-to-Event Endpoints*. Chapman and Hall, 2009.

J. Robins, L. Orallana, and A. Rotnitzky. Estimaton and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*, 27(23): 4678–4721, 2008.

J. M. Robins and A. Rotnitzky. Comment on the Bickel and Kwon article, "Inference for semiparametric models: Some questions and an answer". *Statistica Sinica*, 11(4):920–936, 2001a.

J. M. Robins, M. A. Hernan, and B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560, 2000a.

J. M. Robins, A. Rotnitzky, and M.J. van der Laan. Comment on "On Profile Likelihood" by S.A. Murphy and A.W. van der Vaart. *Journal of the American Statistical Association – Theory and Methods*, 450:431–435, 2000b.

J.M. Robins. Robust estimation in sequentially ignorable missing data and causal inference models. In *Proceedings of the American Statistical Association*, 2000a.

J.M. Robins. Discussion of "Optimal dynamic treatment regimes" by Susan A. Murphy. *Journal of the Royal Statistical Society: Series B*, 65(2):355–366, 2003.

J.M. Robins. Optimal structural nested models for optimal sequential decisions. In Heagerty P.J Lin, D.Y, editor, *Proceedings of the 2nd Seattle symposium in biostatistics*, volume 179, pages 189–326, 2005a.

J.M. Robins. Optimal structural nested models for optimal sequential decisions. Technical report, Department of Biostatistics, Havard University, 2005b.

J.M. Robins. A new approach to causal inference in mortality studies with sustained exposure periods - application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7:1393–1512, 1986.

J.M. Robins. The analysis of randomized and non-randomized aids treatment trials using a new approach in causal inference in longitudinal studies. In L. Sechrest, H. Freeman, and A. Mulley, editors, *Health Service Methodology: A Focus on AIDS*, pages 113–159. U.S. Public Health Service, National Center for Health SErvices Research, Washington D.C., 1989.

J.M. Robins. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proceeding of the Biopharmaceutical section*, pages 24–33. American Statistical Association, 1993.

J.M. Robins. Causal inference from complex longitudinal data. In Editor M. Berkane, editor, *Latent Variable Modeling and Applications to Causality*, pages 69–117. Springer Verlag, New York, 1997a.

J.M. Robins. Structural nested failure time models. In P. Armitage, T. Colton, P.K. Andersen, and N. Keiding, editors, *The Encyclopedia of Biostatistics*. John Wiley and Sons, Chichester, UK, 1997b.

J.M. Robins. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical models in epidemiology, the environment, and clinical trials (Minneapolis, MN, 1997)*, pages 95–133. Springer, New York, 2000b.

J.M Robins and A. Rotnitzky. Comment on Inference for semiparametric models: some questions and an answer, by Bickel, P.J. and Kwon, J. *Statistica Sinica*, 11:920–935, 2001b.

J.M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology*, Methodological issues. Bikhäuser, 1992.

S. Rose and M.J. van der Laan. Simple optimal weighting of cases and controls in case-control studies. *The International Journal of Biostatistics*, page http://www.bepress.com/ijb/vol4/iss1/19/., 2008.

S. Rose and M.J. van der Laan. Why match? investigating matched case-control study designs with causal effect estimation. *The International Journal of Biostatistics*, page http://www.bepress.com/ijb/vol5/iss1/1/., 2009.

P.R. Rosenbaum and D.B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55, 1983.

M. Rosenblum, S.G. Deeks, M.J. van der Laan, and D.R. Bangsberg. The risk of virologic failure decreases with duration of hiv suppression, at greater than 50% adherence to antiretroviral therapy. *PLoS ONE*, 4(9): e7196.doi:10.1371/journal.pone.0007196, 2009.

D.B. Rubin. *Matched Sampling for Causal Effects.* Cambridge University Press, Cambridge, MA, 2006.

D.B. Rubin and M.J. van der Laan. Empirical efficiency maximization: Improved locally efficient covariate adjustment in randomized experiments and survival analysis. *The International Journal of Biostatistics*, Vol. 4, Iss. 1, Article 5, 2008.

A.J. Rush, M. Trivedi, and M. Fava. Depression. *IV. American Journal of Psychiatry*, 160(2):237, 2003.

L. Schneider, Tariot.P.N., K. Dagerman, S. Davis, J. Hsiao, S. Ismail, B. Lebowitz, C. Lyketsos, Ryan J., Stroup T., Sultzer D., D. Weintraub, and J. Lieberman. E?ectiveness of atypical antipsychotic drugs in patients with alzheimers disease. *New England Journal of Medicine*, 355(15): 15251538, 2006.

J.S. Sekhon. Multivariate and propensity score matching software with automated balance optimization: The matching package for R. *Journal of Statistical Sotware, Forthcoming*, 2008.

S. Sinisi and M.J. van der Laan. The deletion/substitution/addition algorithm in loss function based estimation: Applications in genomics. *Journal of Statistical Methods in Molecular Biology*, 3(1), 2004.

M. Swartz, D. Perkins, T. Stroup, S. Davis, G. Capuano, R. Rosenheck, F. Reimherr, M. McGee, R. Keefe, J. McEvoy, J. Hsiao, and J. Lieberman. E?ects of antipsychotic medications on psychosocial functioning in patients

with chronic shizophrenia: Findings from the nimh catie study. *American Journal of Psychiatry*, 164:428436, 2007.

P. Thall, R. Millikan, and H.-G. Sung. Evaluating multiple treatment courses in clinical trials. *Statistics in Medicine*, 19:10111028, 2000.

C. Tuglus and M.J. van der Laan. Targeted methods for biomarker discovery, the search for a standard. *UC Berkeley Working Paper Series*, page http://www.bepress.com/ucbbiostat/paper233/., 2008.

M.J. van der Laan. Causal effect models for intention to treat and realistic individualized treatment rules. Technical report 203, Division of Biostatistics, University of California, Berkeley, 2006.

M.J. van der Laan. Estimation based on case-control designs with known prevalance probability. *The International Journal of Biostatistics*, page http://www.bepress.com/ijb/vol4/iss1/17/, 2008.

M.J. van der Laan and S. Dudoit. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: Finite sample oracle inequalities and examples. Technical report, Division of Biostatistics, University of California, Berkeley, November 2003.

M.J. van der Laan and S. Gruber. Collaborative double robust penalized targeted maximum likelihood estimation. *The International Journal of Biostatistics*, 2009.

M.J. van der Laan and M.L. Petersen. Causal effect models for realistic individualized treatment and intention to treat rules. *International Journal of Biostatistics*, 3(1), 2007.

M.J. van der Laan and J.M. Robins. *Unified methods for censored longitudinal data and causality*. Springer, New York, 2003.

M.J. van der Laan and D. Rubin. Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1), 2006.

M.J. van der Laan, S. Dudoit, and S. Keles. Asymptotic optimality of likelihood-based cross-validation. *Statistical Applications in Genetics and Molecular Biology*, 3, 2004.

M.J. van der Laan, M.L. Petersen, and M.M. Joffe. History-adjusted marginal structural models and statically-optimal dynamic treatment regimens. *The International Journal of Biostatistics*, 1(1):10–20, 2005.

M.J. van der Laan, S. Dudoit, and A.W. van der Vaart. The cross-validated adaptive epsilon-net estimator. *Statistics and Decisions*, 24(3):373–395, 2006.

M.J. van der Laan, E. Polley, and A. Hubbard. Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6(25), 2007. ISSN 1.

M.J. van der Laan, S. Rose, and S. Gruber. Readings on targeted maximum likelihood estimation. *Technical report, working paper series http://www.bepress.com/ucbbiostat/paper254/*, September, 2009.

A.W. van der Vaart, S. Dudoit, and M.J. van der Laan. Oracle inequalities for multi-fold cross-validation. *Statistics and Decisions*, 24(3):351–371, 2006.

Z. Yu and M.J. van der Laan. Construction of counterfactuals and the G-computation formula. Technical report, Division of Biostatistics, University of California, Berkeley, 2002.

Z. Yu and M.J. van der Laan. Double robust estimation in longitudinal marginal structural models. Technical report, Division of Biostatistics, University of California, Berkeley, 2003.